# LEXICAL OPPOSITION

## An investigation into canonical and peripheral phenomena of antonymy


Sandra Kotzor

Kellogg College

University of Oxford


A thesis submitted for the degree of

*Doctor of Philosophy*

Trinity Term 2010

# Acknowledgements

This thesis would not be in its present form without the help and encouragement of a large number of people. First and foremost, I would like to thank both my supervisors (in chronological order): Dr. David Cram and Prof. Aditi Lahiri. From the very beginning of my time at Oxford, David has shown great interest in my work and has been extremely generous with his time. His support and sharp intellectual insights have been essential to my work. Aditi's help with the experimental component of my thesis has been invaluable as has her guidance and constructive feedback during the latter stages of its completion.

A variety of other people have contributed academically to different aspects of my work. I would like to thank Henning Reetz and Adam Roberts for their help with my experiments and, in particular, with the statistical analysis of my data. I am also very grateful for the detailed and very helpful feedback I received from John Coleman, Nivedita Mani and Mary Dalrymple at various stages of my doctoral research.

I have had the opportunity to present my work in various fora and many of the comments and discussions which grew out of these presentations have helped shape my research. Particular thanks are due to the participants of the Modern Linguistics Seminar at the LMU Munich, especially to Dr. Wolfgang Falkner, Prof. Hans-Jörg Schmid and Prof. Richard Janney for their perceptive comments and continued support.

The completion of this thesis would not have been possible without the generous financial support of the FAZIT Stiftung and the opportunity to work as Junior Dean at Kellogg College. Their help is gratefully acknowledged.

I have also received much (non-academic) support from many other sources which has been crucial during my doctorate and I would like to express my gratitude here. Firstly, to the Linguists for weekly pub nights, lunches and many a long night in the lab, and secondly, to my non-linguist friends back in Munich and here in Oxford who have (mostly) managed to keep a healthy distance from my work and have thus been invaluable in keeping me sane. Some individuals from the above groups deserve a special mention for providing unfailing support, laughs, food, tea and proofreading: Alex, Allison, Amy, Ben, Bozhil, Ellie, Holly, Liz, Sam and Kate.

There are two more people who have been instrumental in the completion of this thesis both in a metaphorical and a very concrete way. They have proofread and commented on earlier and final drafts, and have always listened patiently. Their support, advice and reassurance were, at times, much needed. Rachele and Steven – thank you so much. I cannot imagine what I would have done without you.

Lastly, and most importantly, I would like to thank my parents and my siblings without whom none of this would have been possible. Thank you for supporting me in every way and making me curious about language while never letting me forget that my work is not the only thing that counts – this thesis is dedicated to you.

## Table of Contents

# Abbreviations and typographical conventions

## Abbreviations

| | |
|---|---|
| ANOVA | Analysis of Variance |
| BNC | British National Corpus |
| EAT | Edinburgh Word Association Thesaurus |
| fMRI | functional Magnetic Resonance Imaging |
| GOE | Goodness-of-exemplar |
| ICM | Idealised Cognitive Model |
| LD | Lexical Decision |
| LDCE | Longman Dictionary of Contemporary English |
| MHG | Middle High German |
| ModG | Modern German |
| OALD | Oxford Advanced Learners' Dictionary |
| OE | Old English |
| OED | Oxford English Dictionary |
| OHG | Old High German |
| RT | Reaction Time |
| SOA | Stimulus Onset Asynchrony |
| St Dev | Standard Deviation |

## Typographical conventions

| | |
|---|---|
| *word:word* | antonym pair |
| <bat> | grapheme |
| *italics* | language example |
| **bold** | terminology where it is first mentioned and defined, emphasis |
| CAPS | categories |
| 'attribute' | attributes / features |
| SMALL CAPS | concepts |
| <u>SMALL CAPS</u> | cognitive models |
| / / | phonetic transcription |
| 'meaning' | meaning of a lexeme |

## Index of Figures

### CHAPTER 2



### CHAPTER 3



### CHAPTER 4

## CHAPTER 5

## CHAPTER 6

## Index of Tables

## CHAPTER 6

# 1. Preliminaries

There's more than one way to be right
About the opposite of *white*,
And those who merely answer *black*
Are very, very single-track.
They make one want to scream, "I beg
Your pardon, but within an egg
(A fact known to the simplest folk)
The opposite of white is *yolk*!"
(Wilbur 2000: 17)

The idea that human thought and the way we see the world are determined by binary concepts is not a recent one; it has been considered by many and there are countless examples which lend support to this view - not just in language[1]. The mere fact that Richard Wilbur chose to write a whole series of children's poems on opposites highlights that these antonymical word pairs form part of a child's 'education' (in the widest sense) and are offered as a structuring mechanism to make sense of the world.

On a more academic note, binary concepts, like *dichotomy* for example, are key terms in Western economics, sociology, mathematics, philosophy, logic and many other fields which deal with a wide range of different topics. From as early as the 4th century BC, opposition has been a key component of logic and Aristotle's square of opposition[2] provided the basis for much research in that field.

From a linguistic perspective, much work has been done on binarity, negation (cf. Zimmer 1964, Horn 1989, Kjellmer 2005) and opposition (cf. Ogden 1967, Clark 1972[3], Lehrer 1985) in fields such as formal semantics, syntax and lexical semantics as well as in computational linguistics and language acquisition. Lakoff and Johnson (1980), for example, provided support for the idea that opposition is a fundamental structuring principle of the way we think by illustrating the construction of a number of conceptual metaphors with a fundamentally binary structure which underlie many of the metaphorical expressions we use in everyday language (e.g. *she is up* vs. *she is feeling down*). In the present research, I will take a closer look at the words we use to express opposed concepts and investigate the internal structure of the category of opposition. This study will furthermore investigate which factors make an opposed pair of words or concepts a 'good'

---

[1] Needham (1987: blurb) calls opposition 'an elementary and necessary mode of thought' and a 'constant feature of the human mind' while Saussure (1916) considers opposition one of the basic tenets of structuralism.
[2] See Kneale & Kneale (1962) for a discussion of Aristotle's square of opposition.
[3] The developmental angle is a very interesting one and has been investigated in several studies aside from Clark 1972 (cf. for example Jones & Murphy 2005 for a corpus approach to antonym acquisition or Nelson 1977 for a discussion of the paradigmatic/syntagmatic shift). It will not be discussed in detail in the present research since the focus of this study is a different one.

opposite pair and indeed attempt to shed more light on the question whether this opposite relation is one between concepts or between the words which encode these concepts.

The Oxford English Dictionary (OED) provides two definitions of *opposition* as it is used in linguistics (OED Online[4]):

(a) The state or condition of being opposite in meaning; the relationship between antonyms; antonymy
(b) A functional, or potentially functional, contrast between linguistic (esp. phonological) elements.

The present work will focus on definition (a) and investigate the relationship between lexemes[5] which encode opposing concepts, to determine why some pairings are considered better than others by speakers of a language as well as the overall organisation of antonyms.

Despite the amount of research which has been conducted in the field of opposition, there are several important questions which remain unanswered. Firstly, ever since the first empirical investigations into opposition, there has been a debate about whether antonymy is a lexical relation between words or a semantic one between the concepts encoded by said lexemes (cf. among others Gross et al. 1989, Miller & Fellbaum 1991, Charles et al. 1994, Murphy 2003). Is antonymy simply based on the contrastive association of certain words which becomes conventionalised in our mental lexicon, or are the meanings of these words, and their oppositeness, the guiding factor in the association process? There is still much discussion on this issue and the choice of methodology based on the multi-method approach used in the present research aims to address this question more fully, in particular through the comparison of an English and a German dataset. The comparison between English and German antonym pairs allows for insightful conclusions about the nature of antonymy in general.

Another key question, which has yet to be answered to a satisfactory standard, is which factors influence speakers' perceptions of antonymic strength. Several criteria have been proposed (cf. Chapter 2.5) but a thorough investigation of these features has not yet been carried out since most previous empirical studies of opposition were either focused on a particular type of antonym (usually gradable adjectival opposites like *hot:cold*), thus neglecting the cases on the periphery of the category, or simply too restrictive in the methodology used. The novelty of the current research is not only the inclusion of a large number of opposites which are

---

[4] All quotations from the OED online were taken from www.oed.com (accessed June 2010).
[5] *Lexeme*/*word* and *word field*/*lexical field* are used interchangeably in this study.

not considered part of the central core of antonyms but also the use of different empirical measures to gather several sets of data which are then compared and contrasted to reveal discrepancies and similarities which provide answers to the questions raised above. Both these questions will be considered in greater detail in 1.3 after a more thorough introduction to the phenomenon of lexical opposition.

## 1.1 What is lexical opposition?

The concept of opposition is, as mentioned above, one with a long theoretical history. There are several terms which are used for the same, or very similar, phenomena. The English words *opposition* and *opposite* are both in regular use from the 14th century onwards (www.oed.com) and, although both have been used in narrower definitions within specialist terminology, they are very firmly part of speakers' everyday vocabulary.

*Antonym*, however, is a more specific term which speakers of English outside the linguistics community are largely unfamiliar with. It is furthermore a comparatively recent addition: C.J. Smith coined the term *antonym* in 1867 (in his work *Synonyms & Antonyms*) as a counterpart to *synonym*:

> The Etymology of the word [antonym] merely expresses the idea of one word in *substitution for*, which in matters of verbal debate, is equivalent practically to *opposition to* another; a double force which, in addition to its analogy to Synonym, seems to render Antonym a preferable word to *Counterterm*.' (Smith 1867: Preface 6)

In dictionaries it is now usually defined as a technical term for *opposite* which is applied only to lexical items:

> OED 1989: a term which is the opposite or antithesis of another, a counter-term
>
> OALD 2006: noun (technical) a word that means the opposite of another word; SYN opposite: 'old' has two possible antonyms: 'young' and 'new'. – compare synonym

In linguistics, the use of the term *antonymy* is heavily influenced by the categorisations of opposition proposed by Lyons 1977 and Cruse 1986 (cf. also Cruse 1976) in which they use *antonymy* solely to refer to gradable adjectival pairs of opposites (e.g. *hot:cold* or *long:short*). In recent studies, however, the term is used in its 'dictionary' meaning and also denotes other phenomena of opposition besides the prototypical one of gradable adjective pairs. This is also the definition of the term *antonymy* I will be using in this study: a concept which subsumes all phenomena of opposition which are perceived as such by the average native speaker of a language.

The concept of lexical opposition and the fact that there are some words which have opposites and others which do not are well known to any speaker of English. If someone in the street was asked for the opposite of *happy*, they would almost certainly respond very quickly with either *unhappy* or *sad* and if asked for the opposite of *long* the response would be *short*. No complicated explanations are needed of what constitutes an opposite – it seems to be an intuitive decision which is made very rapidly. However, the question *What is the opposite of* work*?* may be more difficult to answer and the responses would be more varied: for example, *fun*, *free time, play* or *relaxing*. Laypersons find it relatively easy to determine the degree of opposition between two concepts or words and are able to make clear distinctions between what they consider 'good' and 'bad' opposite pairs. There are many cases where this decision depends very strongly on the context in which these concepts are presented since there are many instances in which the opposition is only clear when it is foregrounded by the context. A good example would be *dry:sweet* which is a very good opposite pair when used in the context of choosing wine or sherry but not an ideal pairing when talking about food. While there are many word pairs which satisfy the criterion of being in some way conceptually opposed (*bow:curtsey* for example – one is done by a man and one by a woman), there are only a fairly small number which are extremely closely related and are the kind of pairings where, when people are asked, their answer can be predicted with almost 100% certainty. This difference is one of the main issues under investigation in the present research.

Opposites are extremely frequent in everyday speech and are used by public speakers as a rhetorical tool (e.g. *we take care of everyone, the young and the old*), for emphasis and to create or underscore a contrast (e.g. '*It was the best of times, it was the worst of times'*, Dickens' *A Tale of Two Cities*). Children are introduced to the concept of opposition very early. We tend to divide things into *good* and *bad*, *allowed* and *forbidden*, *right* and *wrong* in an attempt to simplify the complex situations with which we are confronted in everyday life. Children seem to accept binary concepts as something entirely natural and are only introduced to the 'grey areas' as they mature. These are, however, not the only contrasts to which children are exposed: many children's books and learning aids contain a large number of opposite pairs and there are even games dedicated solely to the recognition of opposites.[6] It has been shown (Jones & Murphy 2005, Jones 2007) that children use opposites as frequently as adults and in a variety of functions (cf. 2.2 & 2.2.2 and Clark 1972) from very early on in their language development and the success of the above-mentioned games and learning aids shows that this binary conditioning comes easily.

This prevalence of opposites, and the ease with which we accept them as a natural part of our linguistic repertoire as well as a tool which aids us in compartmentalising our experiences, suggests that the concept of opposition is indeed one of the most fundamental cognitive structuring principles.

There is also evidence that lexemes which encode opposing concepts are related closely to each other in the mental lexicon. This evidence comes from two investigative strands: psycholinguistic experiments such as elicitation tasks, word association and lexical decision tasks, and the analysis of speech errors. The study of mistakes made in natural speech provides insights into how utterances are constructed and at which points this production can go wrong (cf. among others Fromkin 1973 & Cutler 1981). Replacing one member of an antonym pair with another is a relatively common selection error and utterances such as *Go and have a* ***cold*** *shower!* (instead of *hot*) or *Well, she's not very* ***young***. (instead of *old*) occur fairly frequently. They are also not always detected as such since they are only noticeable when the meaning is at odds with the context in which they occur, as the sentences above are grammatically and semantically correct. The linguistic analysis of selection errors is difficult since a very large amount of data would have to be collected, listened to and transcribed to amass a sufficiently large corpus of these mistakes; therefore this method is not used in the present research. The evidence collected by psycholinguistic experiments, for example word association (among many others Deese 1964/1965), elicitation (Paradis et al. 2009) and lexical decision tasks (Gross et al. 1989, Charles et al. 1994, Sabourin & Libben 2000), will be discussed in Chapters 2 and 4.

Recently, research on antonyms has experienced a revival and there are several scholars currently involved in the study of opposition. When this project was conceived in 2006, multi-method research to lexical opposition had not been carried out (cf. 6.3; discussion of Paradis et al. 2009). Many of the early studies (e.g. Lyons 1977 and Cruse 1986, Cruse & Togia 1995) were introspective and did not involve any empirical methods, while many others were corpus-based (Hoey 1991,

---

[6] For example *Learning to read. Antonyms (Opposite Words)* by Creatives (Creative Educational Aids PVT. LTD.)

Mettinger 1994, Fellbaum 1995 and Jones 2002) or involved judgement tasks or lexical decision tasks (Gross et al. 1989, Murphy & Andrew 1993, Charles et al. 1994).

There have also been approaches which have considered the treatment of antonyms in dictionaries (e.g. Paradis & Willners 2007), while Müller (1998) has compiled an antonym dictionary (Gegenwort-Wörterbuch) in German which contains not simply the opposite pairings one would find in a generic thesaurus but goes far beyond that by including pairs such as *Knicks:Diener*[7] ('curtsey:bow'), *Handarbeit:Maschinenarbeit* ('made by hand' vs. 'made by a machine') or *Froschperspektive:Vogelperspektive* ('worm's eye view' vs. 'bird's eye view') and many even more highly contextualised examples. Research in the field of antonymy has been moving away from the traditional approach of introspective and textual analysis of gradable adjectival antonyms. The most recent treatments of antonymy have begun to incorporate a variety of empirical methods (cf. Paradis et al. 2006), and one study in particular (Paradis et al. 2009) takes a similar approach to that used in the present research which is outlined in the following section. Their work will be discussed in detail alongside a summary of my results in the concluding chapter of this thesis (cf. 6.3).

## 1.2 The current approach[8]

The approach to opposition applied in this study is more inclusive than previous ones in order to include an investigation of the boundaries of the category, since it is very frequently the case that an analysis of the differences between peripheral and central cases of a phenomenon lead to valuable insights as to what makes the central members particularly good representatives (cf. prototype categories – Taylor 1989 & 1990, Vandeloise 1990, Cruse 1990 & 1994, Ungerer & Schmid 2006). Therefore, the scope of opposite pairs included in the study is wider than in previous studies and includes pairs from different word classes and a large number of non-canonical pairs (e.g. *excellent:atrocious*, *work:play*, *king:queen*, *buy:sell*), some of which are deliberately modified to generate weaker responses than related

[7] The convention *lexeme:lexeme* will be used for all opposite pairs in the present research.
[8] The approach in this study is a synchronic one and while, in some individual cases, the diachronic perspective is touched upon, it is not considered in its own right (cf. also 7.4).

canonical pairs (e.g. *freezing:hot* vs. *cold:hot*) to allow for systematic investigation of the factors which influence the judgement of the strength of an antonymic relation.

Furthermore, the present research makes use of several distinct empirical methods to investigate antonymy: corpus studies, judgement tasks and lexical decision experiments. The results of these three empirical techniques all bring their individual contribution to the research and the combination of the data collected throughout this study is not only more comprehensive but allows for a precise analysis of the discrepancies between the three data sets. These differences as well as the patterns the data shows are of crucial importance to the theoretical questions this research is aiming to answer. The selection of the word pairs under investigation in addition to the multi-method approach is what makes this project unique in antonym research and promises conclusions which will not only advance the study of antonymic relations in lexical semantics and psycholinguistics but will also allow assessment of the adequacy of certain investigative techniques.

## 1.3 Research question(s)

This study aims to answer three overarching questions which are dependent on each other: the answer to the first question determines the scope of the second and both have an impact on the treatment of the third.

(1) Is antonymy a gradient phenomenon or is there a clear distinction between canonical and non-canonical opposites?

(2) Which factors determine where on the antonym continuum a lexeme will be placed or, if the data supports a division between canonical and non-canonical opposition, how are word pairs assigned to one or the other category?

(3) Is antonymy a relation between concepts or between the lexemes which encode these concepts? How are antonyms stored in the mental lexicon?

It has been assumed by many researchers (among others Cruse 1986, Lyons 1977, Gross & Miller 1990, Mettinger 1994, Muehleisen 1997, Jones 2002 and Murphy 1994/2003)[9] that antonyms can be divided into two clearly defined sub-groups: a category of 'canonical' opposites and one of 'non-canonical' pairs. However, it is never made altogether clear which features a word pair has to possess to qualify for membership in the former category and whether it automatically falls into the latter,

[9] This is not to say that all these scholars support the 'lexical categorical model' (Gross & Miller 1990) but they have worked with the assumption that there are two groups of opposites which are, at least to some extent, definable.

if it does not meet a certain set of criteria. Questions (1) and (2) are both related to this previous theoretical standpoint. Recently, there have been approaches which are more heavily based on empirical data and which show that there may be grounds to support a continuum approach to antonymy. However, the question of which criteria influence the 'goodness' of any given opposite pair remains. The search for factors which allow for an accurate prediction of antonymic strength of any opposite pair is a key component of this study. The data collected using the three methods introduced above will shed light on whether a continuum approach can be supported or whether there is a clear distinction between central and more peripheral members of the category of lexical opposition.

Question (3) is the focal point of a long-standing debate spanning at least two decades, with one camp advocating a conceptual approach and claiming that the lexical relatedness and high associative strength of conceptually opposed word pairs is a result of this opposition and the other claiming that the strong associative bond between the two lexemes, which is entrenched through frequent co-occurrence, is the source of their high degree of opposition. Both sides have provided evidence which supports their arguments but this evidence is rarely comprehensive and often includes only very central phenomena of opposition.

Several other hypotheses will be proposed throughout the data analysis chapters to explain certain patterns observed in the data. These will be introduced alongside the relevant data and tested and their verification will support the argument in the last chapter when the discussion returns to the three central questions above.

## 1.4 Outline of thesis

The present work contains two theoretical chapters (Chapters 2 and 6), which bracket three chapters reporting on the results yielded by the data collected, of which each presents a different empirical approach. Chapter 3 presents the analysis of data gathered in a corpus study of the co-occurrence of the members of 210 word pairs in comparison to the results of a judgement task on the same word pairs. In Chapter 4 a behavioural lexical decision task forms the basis for discussion, while Chapter 5 adopts a contrastive approach, comparing the English data presented in the previous chapters with German opposite pairs. Chapter 2 gives a general overview of previous research on antonymy and raises a number of questions regarding the hypotheses put forward above, while Chapter 6 takes stock of what has been accomplished regarding those questions and which questions highlighted in Chapter 2 the multi-method empirical investigation has managed to answer. Each chapter will briefly be outlined below.

### CHAPTER 2

The aim of this chapter is to provide a framework for the current research by presenting an overview of previous research on antonymy from various perspectives. The different approaches – theoretical, introspective, corpus-based and experimental – will be presented and discussed in turn to provide the theoretical basis for much of the discussion and the data analysis presented in the following chapters, as well as to illustrate the necessity of the present research by putting it into context.

A small number of psycholinguistic investigations into antonymy have been reported on (among others Gross et al. 1989, Charles et al. 1994, Hermann & Chaffin 1986, Sabourin & Libben 2000 and, most recently, Paradis et al. 2009) and are included in the overview in Chapter 2 with regard to their theoretical merit but, since some of methodology is similar to that used in the present research, the experimental details are only discussed thoroughly in Chapter 4. Some neurolinguistic research (cf. Gazzaniga & Miller 1989, Varley 1991 and Jeon et al. 2009), as well as some theoretical constructs directly related to the methodology discussed in Chapter 4, are also introduced only at this later stage (cf. CHAPTER 4 below).

The summary of previous research concludes with a section which investigates the question of what canonical opposition is and how we can attempt to define such a concept. All those criteria which have been proposed for 'good' opposition in the research discussed are gathered together, presented and illustrated using the example of the opposite pairs on the TEMPERATURE scale. This anticipates some of the discussion in Chapter 3 which will determine how much explanatory power these criteria have for the data collected in the judgement task.

### CHAPTER 3

This is the first part of the empirical component of the study. The main aim of this section is to investigate the relationship between associative strength and antonymic strength and to determine how certain factors influence the latter.

Associative strength is measured in two ways for the purpose of the discussion in this chapter: frequency of co-occurrence in the British National Corpus

(BNC) and number of elicitations in the Edinburgh Word Association Thesaurus (EAT). These two measures are compared to the data collected in a goodness-of-exemplar (GOE) rating task where 160 subjects rated 210 word pairs on a scale from 1 (excellent opposite pair) to 7 (very poor opposite pair). This is taken as an indication of each pair's antonymic strength. The end of the chapter also contains data from another source, an attribute listing task, to illustrate the category-internal structure of certain concepts.

After an explanation of the methodology used to collect the corpus and judgement task data, a number of clusters of different types of antonymy are considered in detail. Their respective measures of antonymic and associative strength are compared, and possible reasons for discrepancies between and differences within the two variables are discussed using the criteria outlined at the end of Chapter 2. The clusters, which serve as examples throughout the present research, fall into three categories:

(a) gradable opposites (TEMPERATURE, MERIT and SIZE)
(b) complementaries (GENDER)
(c) converses (verbal and nominal)

Conclusions are drawn based on the GOE-rating data as to which factors seem to influence the degree of antonymic strength of an opposite pair most. Some new hypotheses are put forward for discrepancies which cannot be explained by any of the previously mentioned criteria and these are put to the test in Chapter 4, where the next set of data is added to the analysis.

### CHAPTER 4

Chapter 4 focuses on the analysis of the behavioural data collected in two experiments, which is added to the results presented in Chapter 3 to further substantiate the claims made about the importance of individual factors as well as to investigate how accurate a picture of lexical representation of opposites the speaker judgements from the GOE-rating task provide.

Before setting out to introduce the methodology and experimental design used in the two studies reported on in this chapter, some central theoretical constructs are introduced and explained since they are the building blocks upon which the experiments are based. Spreading activation, a very brief introduction to models of lexical access and recent research on the representation of lexical opposition in the mental lexicon are discussed and evaluated from both a psycholinguistic and neurolinguistic perspective.

After an explanation of the statistical methods used, the two experiments are analysed separately with Experiment 1, a slightly modified lexical decision task, constituting the bulk of the discussion. It was designed to investigate the discrepancies observed in the data in Chapter 3 more closely and to see whether the decisions made under time constraints in a lexical decision task match those from the judgement task. Experiment 2, a much shorter lexical decision task with priming, is discussed in the last section of Chapter 4 before a summary of the discussion.

CHAPTER 5

The last of the data chapters introduces the German data gathered from translated versions of the GOE-rating task and the behavioural study (Experiment 1). The judgement task data and behavioural results are discussed separately.

The first section gives a brief overview of the German GOE-rating data in comparison with the associative measures (t-scores) calculated from the frequency of co-occurrence data which was collected from the biggest German language corpus (IDS Mannheim) with the COSMAS concordancer. After this summary, cases which display noteworthy discrepancies between the English and German results are discussed individually in similar groups as in the English analysis above (converses, scalar antonyms, complementary opposition (gender) and a case of four overlapping adjectival opposites which displays structural similarities to that of the English size adjectives).

The second part of the chapter is dedicated to the analysis of the German behavioural data. In analogy to the treatment of the GOE-rating data above, a short summary of the results is given to determine whether the same overall patterns are observed in both languages. This is followed by a discussion of individual cases where pairs which displayed differences in the GOE-rating are investigated and discussed in light of the additional behavioural results and the reasons for the discrepancies are re-evaluated. The analysis of noteworthy differences which stem only from the behavioural data concludes this chapter and with it the presentation and analysis of the collated data.

## CHAPTER 6

The aim of the last chapter before the concluding remarks is to bring the focus back to theoretical considerations which have been affected by the data and to re-evaluate the hypotheses proposed throughout the data chapters to explain individual and occasionally even unique phenomena. Furthermore, the broader research questions formulated in 1.3 and 2.5 are revisited and answers are proposed.

Two main questions lie at the centre of the discussion: the influence of individual factors on the antonymic strength of word pairs, and the larger question of whether antonymy is a gradient phenomenon with the pairs patterning along a continuum from the canonical antonyms to pairs which are barely antonymic and can only be considered opposites in certain contexts.

The former question is answered by considering each of the main factors in turn and evaluating the evidence collected and analysed in the data chapters. The discussion includes, besides the factors listed in 2.5, an analysis of the influence of minimal difference, sequencing and antonym type on antonymic strength. A holistic approach to antonym canonicity is proposed to conclude this section. The latter question, which is connected to the debate whether antonymy should be seen as a lexical or semantic relation and how the category of opposition is structured internally, is answered in detail in 6.3 where the results found in this study are compared to the recent multi-method approach by Paradis et al. (2009).

## CHAPTER 7

The concluding remarks discuss the benefits of a multi-method approach including peripheral examples of antonymy, which enables a more wide-ranging and thorough analysis of the complete category of opposition than the more common narrower design which focuses heavily on adjectival antonym pairs. Improvements on the methods used in the study are proposed and avenues for further investigation are highlighted, before the concept of antonymy is re-defined on the basis of the theoretical and experimental findings presented throughout the present research and the definition given in Chapter 2.2.2 is re-visited and altered to incorporate those findings.

# 2. An overview of perspectives on opposition

Because what's *present* doesn't last,
The opposite of it is *past*.
Or if you choose to look ahead,
*Future*'s the opposite instead.
Or look around to see what's here,
And *absent* things will not appear.
(Wilbur 2000: 19)

Almost all accounts of antonymy (among others Lyons 1977, Cruse 1986, Mettinger 1994, Muehleisen 1997, Jones 2002 and Murphy 1994/2003) have in one form or another made use of the term, or at least the concept of, **canonical antonymy**. This term is understood to mean that there are antonyms which belong to a certain central group, for instance *happy:sad* or *big:small*. However, in all these works on antonymy this term is taken to be self-explanatory and is never explicitly defined. In this chapter, I will be looking at various approaches to antonymy put forward by different scholars and will attempt to factor out the characteristics of canonical antonyms proposed in these studies. Furthermore, I will assess their validity and practical applicability.

First of all, I will look at the broader category of opposition and introduce several classifications which have been central to the research conducted in this field and discuss relevant approaches from different disciplines, before looking at the notion of a 'canonical antonym' in a little more detail (cf. 2.5). Each possible feature of canonical antonyms will be discussed and illustrated in turn in order to assess its importance and explanatory power for the centrality of canonical antonyms. An important part of this discussion will be the attempt to find an answer to the question whether antonymy should be considered a relation between lexical items or a relation between concepts (cf. 2.5 & 6). This distinction has led to considerable discussion but a satisfactory answer has not yet been found.

## 2.1 Classifications of opposition

Most studies have mainly looked at one type of antonym: the gradable, adjectival antonym (*good:bad*), as opposite pairs of this type display the most overtly interesting characteristics (such as markedness and polarity) with regard to their behaviour in, for example, question formation, comparatives and superlatives.

Adjectives are generally considered to be the prototypical word class for antonymic relations. However, there are also numerous well-established nominal and verbal antonym pairs which are not derived from, or related to, an adjectival pair. Hence a comprehensive discussion of antonymy has to take into account that, despite being most prominent and prototypical among adjectives, the phenomenon also occurs in all other open word classes and even, very strongly, among certain prepositions (cf. Cruse 1986).

Probably the best-known discussions of categories of antonymy are those presented by Lyons (1977) and Cruse (1986). Although their introspective approaches have been criticised by recent scholars as subjective and not empirically substantiated, they have nevertheless laid the groundwork in this field and have provided many linguists with a basis from which to conduct their own investigations of antonymy. Furthermore, the categories they proposed (cf. Figure 2.1) have given important insights into the behaviour of antonyms. Both have also made attempts at describing the features of the canonical antonym and it is from their studies that we gain some of the most basic features listed below (cf. 2.5). This makes these investigations useful even though they are ‘entirely dependent on the intuitions of their proponents’ (Jones 2002: 20). Therefore, a discussion of these categorisations in some detail is essential. They also contribute to the illustration of the chronological development of the study of antonymy.

### 2.1.1 Lyons 1977

John Lyons, who bases his work on antonymy on the categories proposed by Edward Sapir,[1] states in his well-known work on semantics (1977) that the ‘most common opposites in English tend to be morphologically unrelated’ but that they are outnumbered by morphologically related ones. He claims that morphologically unrelated antonyms show the dichotomy involved in antonymy best as the words have nothing at all in common (e.g. *good:bad*, *high:low*).

Lyons proposed subdividing the class of opposites into two distinct subclasses with several additional subdivisions (cf. Figure 2.1). His initial distinction was drawn between **binary** and **non-binary** contrasts. I will first of all deal with the non-binary contrasts as these are more easily explained. Here, a line is drawn between **serial** and **cyclical** (e.g. days of the week) contrasts: serial contrast is further subdivided into **scales** (e.g. school grades) and **ranks** (e.g. military ranks). Thus, *officer* and *lieutenant* are in a relation of scalar contrast but they would not readily be called opposites by native speakers of English and would presumably not,

[1] This will not be discussed in detail here as this original classification is subsumed in Lyons' more detailed one. For a detailed discussion see Sapir 1944 and Lyons 1977.

or very infrequently, elicit each other in unrestricted elicitation experiments. These non-binary contrasts were excluded from the discussion of opposition in many other studies and I will also not be focussing on these instances here as, in my opinion, opposition is fundamentally binary.

Lyons' classification of binary contrast consists of five categories which I will briefly mention and illustrate, only one of which is assigned the label 'antonymy'.[2] Orthogonal and antipodal opposition are concepts which do not refer to many antonym pairs as they are very rigidly defined. **Orthogonal opposition** (e.g. *north:east*) is the opposition of concepts which are not directly (spatially) opposed to each other but are, as in the example, neighbours on the compass. Similarly, **antipodal opposition** is the opposition of concepts which are spatially directly opposed (*north pole:south pole; north:south*). These two concepts are not strictly speaking relevant to the following discussion since they only include a small number of opposite pairs, although antipodal opposition is included in the category of directional opposition (albeit as a special case).

**Directional opposition** (e.g. *up:down, come:go, left:right*) is defined in spatial terms only and is, in my opinion, one of the most fundamental and most widespread types of antonymy, as direction is often expressed in opposition and the distinction is also more clearly visible than in other, more evaluative antonym pairs, which depend very much on interpretation (*good:bad, clean:dirty*).

Lyons' category of converseness is one I will be addressing in greater detail in the second part of this study. Here, I will therefore only mention that he described converseness as a reciprocal relation, like the one in *buy:sell* or *parent:child*. These relations were also not treated as cases of opposition proper as they cannot be put into one of the categories below, which, according to Lyons (and also Cruse 1986) exemplify the categories of opposition proper.

The distinction made here is that between **gradable** and **non-gradable** opposites. **Antonymy**, in the narrow sense, is gradable and there is a possibility other than the two antonyms which could apply, for example, in the case of *hot:cold*, where *warm* or *tepid* is also a valid concept (*tertium datur*). The scope of the two antonyms does not cover the whole range of possibilities in the relevant field.

(1.1) *The tea was hot. – The tea was not cold.*

(1.2) *The tea was not cold. ≠ The tea was hot.*

[2] A more detailed discussion of Lyons' categories can be found in Cruse 1986 and Mettinger 1994.

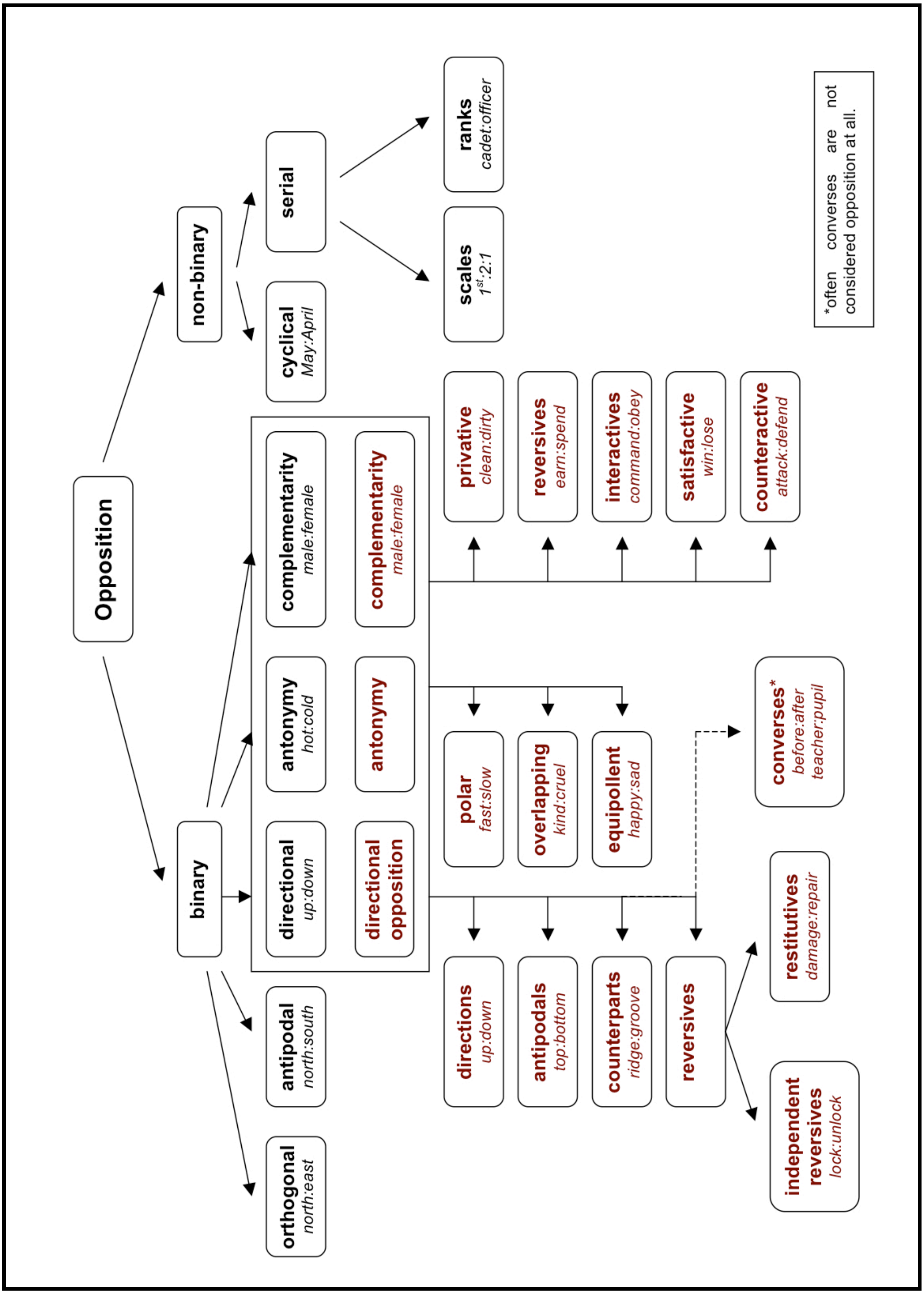



**Figure 2.1** *Classifications of antonymy (Lyons 1977 & Cruse 1986)*

In the second case, that of **complementarity**, they do precisely that. There is no third possibility (*tertium non datur*), for example in *married:unmarried*, *dead:alive* or *male:female*. These antonym pairs exhaust the relevant category and the negation of one implies the affirmation of the other.

(1.3) *The budgie was dead. – The budgie was not alive*.
(1.4) *The budgie was alive. – The budgie was not dead*.

This distinction between gradable and non-gradable antonymy is an important one as far as truth relations are concerned and is also evident in the fact that non-gradable opposites cannot generally be combined with quantifying adjectives (*very, quite* etc.) unless in a certain pragmatic context when a scale of the properties encapsulated by the antonym pair is suggested (e.g. *I am very much alive*). In this case the lexeme *alive* has a slightly different semantic content than in its usual, non-gradable sense.

It is not clear whether the distinction between gradable and non-gradable opposition is perceived as a relevant one by speakers of the language (cf. also Chapters 4 & 6) and whether, apart from the usage restrictions with quantifiers, it makes a difference in use or in storage in the mental lexicon. However, syntactically it is a fundamental distinction, as gradable and non-gradable opposites show different properties which are relatively easily captured and distinguished.

### 2.1.2 Cruse 1986[3]

Cruse's extremely detailed and highly theoretical discussion (cf. also Cruse 1976) includes a wealth of examples and counterexamples for every category presented, and gives detailed explanations of the individual phenomena involved. Some of the classifications seem at first rather arbitrary and unnecessarily detailed - which they certainly are from a use-oriented perspective. It is, however, possible that some of the phenomena discussed by Cruse influence the perception of antonymic relations in use and are a deciding factor in the treatment of differences in those relations. Thus a reasonably detailed account of Cruse's classification seems appropriate here, as these categories are also taken into account in the design of the behavioural experiments and judgement tasks presented in Chapters 3, 4 and 5.

Cruse based his classification on Lyons' initial division of opposition and refined it to include various subcategories for the main classes of antonyms. Cruse

[3] For obvious reasons, this section draws heavily on Cruse 1986 (Chapters 9-11) and all information and ideas given, my comments aside, are taken from this work.

divided his discussion of opposition into three parts: complementaries and antonyms, directional opposition, and general questions on opposition.

The first part deals with the most clear-cut cases of opposition, namely those which are considered gradable and non-gradable antonyms by Lyons. Cruse first considers **complementaries** (non-gradable opposites), which he claims are conceptually the simplest cases of opposition as they are exhaustive of one conceptual domain, mutually exclusive, not usually gradable (or modifiable by adverbs such as *extremely* or *slightly*) and often odd in the comparative or superlative. They are usually adjectives or verbs and generally one member of the pair is more prone to gradability.

A special case in this category are adjective pairs like *clean:dirty* and *safe:dangerous*, as these do not conform to the regular characteristics of non-gradable antonyms. These, so-called **privative antonyms**, seem to have some special status at the borderline between gradable and non-gradable opposition and are not usually classified as complementaries as they do not behave like the pairs in this class in the usual tests. To illustrate this by only one example, consider the following sentences:

(1.5) *It is neither clean nor dirty*.
(1.6) *He neither succeeded nor failed.*
(1.7) *? It is neither clean, nor even slightly dirty.* (Cruse 1986: 203)

*Not clean*, does not, as can be seen in example (1.5), entail that something is dirty but it does entail that something is not completely clean and therefore slightly dirty which is why the sentence in (1.7) is usually considered slightly paradoxical. Furthermore, both *clean:dirty* and *safe:dangerous* can be used with a range of modifying adverbs such as *very, fairly, slightly*, and *extremely*, as well as in the comparative. This suggests that, even though these pairs are logically considered complementaries, conceptually they are, at least to some degree, considered gradable.[4]

Cruse further subdivides the category of verbal complementaries into four sub-categories:

- **reversives** (*earn:save:spend*)
- **interactives** (*command:obey:disobey*)
- **satisfactives** (*compete:win:lose*)
- **counteractive** (*attack:defend:submit*)

[4] In the coding for the behavioural task in Chapter 4, these adjectives are grouped together with the complementaries to minimise the groups for coding since they work in similar ways.

The first verb is said to ‘set the scene’ for the other two: if you compete you either win or lose, or if you receive a command you either obey or disobey. There is also a slightly uneven relationship between the opposite pairs formed between these triadic relations as *command:obey* are considered good opposites, as well as *obey:disobey*. However, *command:disobey* is, although logically equally possible, not usually considered a good example of opposition. *Command* will usually elicit *obey* as an antonym whereas *obey* will usually elicit *disobey* as this is considered a better pair than *obey:command.*

Cruse then moves on to a discussion of **antonymy**, a term which he reserves for a very specific class of opposition only, namely that of fully gradable adjectives which denote a certain degree of a variable property. The more strongly intensified these adjectives are, the further apart on a scale they move (*hot:cold* – *extremely hot:extremely cold*). Unlike non-gradable opposites discussed above, antonyms do not strictly bisect a domain; there is a middle region which is, a few exceptions aside (*tepid, warm*), not lexicalised. There also seems to be an implicit norm with one end of the scale moving towards zero and the other towards infinity.

The category of antonyms was then further subdivided into three distinct categories on the basis of the following properties:

a) comparison
b) evaluative polarity
c) committedness (*How X is it*)?
d) scales underlying the concepts (single scale vs. two scales)[5]

The category of **polar antonyms** (*heavy:light, fast:slow*) typically contains antonym pairs which refer to a property which can be measured and is objectively descriptive. There is a pseudo-comparative to each member. These antonyms operate on a scale which is always specific to a certain category and there is no absolute value attached to the concepts. For example, *a heavy bag* will usually still be lighter than *a light elephant*, which illustrates that the value of these antonyms always has to be interpreted in relation to the referent. Cruse calls this a relative scale. Usually only one of the members of a polar antonym pair will form a neutral, impartial question (*How heavy is it?*) whereas the other is considered unusual and strongly marked in a question (*How light is it?*).[6]

**Overlapping antonyms** (*good:bad; kind:cruel*) have evaluative polarity and are characterised by the existence of a pseudo-comparative to one and a true comparative to the other. With regards to polarity ‘one member is usually

[5] This will be dealt with in greater detail in 2.3.1
[6] For a discussion of markedness, see Lehrer 1985 and 2002.

commendatory (good, kind) and the other is deprecatory (bad, cruel)' (Cruse 1986: 208). These evaluations seem to be 'standardised' by cultural and conventional means and are not subjective judgements as these are treated in a different category. Here, Cruse also stipulates **privative antonyms** (*clean:dirty; safe:dangerous*); their special quality has already been discussed under 3.1.1, as a subgroup. Both members of these pairs form normal questions, although one will be committed, one impartial (*How bad is it?* vs. *How good is it?*).

The third group, termed **equipollent antonyms** (*sweet:sour; happy:sad*), is characterised by each member having a true comparative and both members being able to form normal questions which will, in both cases, be committed. The antonyms in this category all describe 'subjective sensations or emotions […] or evaluations based on subjective reactions rather than on "objective" standards […]' (Cruse 1986: 208).

In his chapter on directional opposition, Cruse divides the observable phenomena into four categories: directions, antipodals, counterparts and reversives. **Directions** are the simplest form of directional opposition and include pairs like *up:down* and *forwards:backwards* which denote pure opposite direction. **Antipodals** are constructed slightly differently, as one term of the antonym pair represents an 'extreme in one direction along some salient axis, and the other term denotes the corresponding extreme in the other direction' (Cruse 1986: 224). Examples of antipodals are pairs like *top:bottom*, *full:empty* and *maximum:minimum*. An interesting sub-category is that of **counterparts** which encompasses pairs like *ridge:groove* or *convex:concave*. Both members of these pairs are characterised by a 'deviation or irregularity in an otherwise uniform surface or shape' (Cruse 1986: 225f.) and the two members are oriented in opposite directions.

The last group of directional opposites is that of **reversives** such as *rise:fall* or *tie:untie*. These opposite pairs differ only in terms of directionality and usually denote motion or change in opposite directions (*rise:fall*) or a change between two determinate states (*tie:untie*). Cruse draws a distinction between independent reversives and restitutives. The main characteristic of the category of **independent reversives** is that there is no need for one member of the pair to indicate the restitution of a former state. This group includes all morphologically marked reversives (*lock:unlock*) and all antonyms which are derivationally related to an adjectival opposite pair (*shorten:lengthen*). **Restitutives**, on the other hand, are, as the term suggests, a category of pairs in which one member denotes the restitution of a former state (*damage:repair*) and one member of the pair will be dependent (*repair*), the other independent (*damage*) as there needs to be damage for repairs to take place.

The last category to be discussed here is one that is often hotly debated: **converses**. In some cases converses are considered to be a relation of opposition and are therefore to be included in the wider category of antonyms, but in others they are excluded. There are several different types of converses but all of them are, according to Cruse, based on a fundamentally spatial notion. However, they are not restricted to this domain and can be extended by analogical or metaphorical processes (*below:above; before:after*). Very often this category is also termed **relational opposition**. Cruse states that lexical converses have to be able to express an asymmetrical relationship between two entities, such as the relationship in the following pairs: *father:son, teacher:pupil*, and *husband:wife* (cf. Cruse 1986: 232). Another group which displays these characteristics is that of the comparatives of polar antonyms. The relationship between pairs like *give:receive* is considered indirect converseness. All types of converses will be considered in greater detail in 3.4.2 and 4.3.1.7, compared to their German counterparts in 5.2.1.2, and finally investigated from a more theoretical perspective in Chapter 6.

After this rather detailed discussion of the classifications of oppositeness from an introspective perspective, I will now briefly outline some recent corpus linguistic approaches to antonymy, one of which, Jones 2002, contains an alternative, text-based classification of antonymy.

## 2.2. The corpus perspective

Early corpus studies (Justeson & Katz 1991, Mettinger 1994, Fellbaum 1995) had to rely on information from comparatively small corpora (by today's standards) whereas the latest corpus study investigating antonymy (Jones 2002 – but cf. also Paradis et al. 2007) uses a 280-million-word corpus comprising the text of all issues of the British broadsheet *The Independent* from 1988 to 1996. Unlike in the approaches discussed above, in corpus approaches to antonymy the main focus of the investigation lies in the usage of antonyms in text and the conclusions which can be drawn from certain uses rather than on a classification of antonymy from a language-immanent perspective. These studies have contributed considerably to a better understanding of the everyday functions antonymy fulfils in texts but have to date been mainly limited to written language. Investigations of antonymy in spoken language are extremely important but researchers are limited by the resources currently available. Nevertheless, some research on antonym use in both child and adult spoken language is now being conducted (Jones & Murphy 2005, Jones 2007 and Paradis et al. 2007) but this research is still at a relatively early stage.[7] I will briefly illustrate some of the main corpus studies and their implications for the current research. Others, mainly Justeson & Katz 1991 and Fellbaum 1995, will be discussed in more detail when I examine the phenomenon of frequency of co-occurrence in 2.5.3.

### 2.2.1 An early corpus-linguistic study (Mettinger 1994)

Mettinger draws attention to the vast number of studies which include antonymy but criticises the fact that most observations on antonymy in these works are based on the author's judgement rather than on linguistic evidence and the antonyms listed as examples are, very often, quite a disparate group even though they seem superficially similar. In Mettinger's eyes empirical, corpus linguistic methods are the best remedy for the shortcomings of previous accounts.

His own research is based on ‘structuralist semantic theory with a mentalistic concept of meaning and a treatment of semantic relations as properties of the language-system’ (Mettinger 1994: 2). In the first part of the study, Mettinger focuses on opposites in context and attempts to determine syntactic frames (much like Fellbaum 1995 and Jones 2002) in which opposites habitually co-occur in his corpus of 43 novels (largely British crime fiction). The 350 pairs of opposites considered in the second part of the study are taken from *Roget's Thesaurus of Words and Phrases* (1972) and are analysed using semantic features and semantic dimensions. Both parts of Mettinger's study discuss and raise questions pertinent to the present research.

The two main questions raised in the earlier part of the study are related to the syntagmatic properties of opposites in English:

> (a) are there any contextual/syntagmatic environments favouring the application of opposites; and, if there are, what are their characteristics?
>
> (b) do the opposites used in such contextual/syntagmatic environments share the same status with regard to the linguistic system of English; and, if they do not, what is the basis on which to distinguish various groups? (Mettinger 1994: 35).

[7] During the time the present research was conducted, several corpus studies on antonymy were carried out, most recently an investigation into cross-linguistic discourse functions (cf. Murphy et al. 2009).

In answer to the first question Mettinger posits 10 syntactic frames (e.g. *X or Y*) in which opposites commonly co-occur and discusses his findings in great detail. In the following chapter he draws a distinction between systemic and non-systemic opposite pairs: pairs which match the criteria drawn up for opposites by various scholars and ones that do not. The latter, non-systemic opposite pairs, constitute adequate opposition in context ('parole') but do not fulfil the criteria to be considered fully-fledged opposites on the level of 'langue' (cf. Saussure 1916). This distinction is relevant for the theoretical discussion of contextual antonyms in Chapter 6.

The second part of Mettinger's study is concerned with the dichotomy of **central** and **peripheral** opposition but only the notion of central opposition is of interest to him. Interestingly, among the phenomena considered ‘peripheral oppositeness of meaning’ (Mettinger 1994: 84) is not only **converseness** but also **directional opposition** which is usually considered relatively central. Mettinger's main aim is to ‘describe and explain the semantic relations obtaining between the members of a pair of systemic opposites in terms of semantic dimension and semantic features, or rather, feature-relations’ (Mettinger 1994: 84). The only types of opposition considered are complementarity, antonymy and gradable complementarity. Important factors in his analysis are gradability and scalarity as well as semantic features. He argues that the features employed will express degrees of different values along certain scales. He puts forward several scales along which opposite pairs can be described depending on whether they are, for example, exhaustive of a scale, relational or evaluative. Every single pair is evaluated and listed with examples which makes the overall study a little dry to read but turns it into a treasure trove of examples.

Mettinger's endpoint proves one possible starting point for the present research, and while this is by no means the only criterion involved in a cognitive construal of opposition it is not one that has been considered in the cognitive research described below (cf. 6.3).

> A treatment of opposites from a cognitive point of view would probably have to assume a scale of increasing amount of encyclopaedic knowledge that is necessary for establishing conceptual integrators and differentiators, starting from a zero-value (characteristic of systemic opposites), becoming more with terminological opposites, and reaching a maximum with encyclopaedic opposites. (Mettinger 1994: 162)

### 2.2.2 A use-oriented classification of antonymy (Jones 2002)

In his corpus study of antonymy, Jones focuses on the use of antonymy in text and develops a new list of categories according to the textual functions fulfilled by an antonymy pair, as was already advocated by Justeson & Katz (1991).

Jones decided on a list of 56 antonym pairs chosen on the basis of intuition of what constitutes a good antonym pair, Deese's list of antonyms (Deese 1964: 347-57) and word class (inclusion of verbs, nouns and adverbs as well as adjectives and the inclusion of both morphologically related and unrelated pairs). Frequency was **not** considered a factor but all antonym pairs in the list are, at least, fairly frequent (*boom:recession* is probably the most infrequent example).

The corpus analysis was then conducted and the sentences used for the study collected by hand, as only sentences in which both members of a pair not only occurred but were also used in an antonymous relationship were of interest to the question. The chosen sentences were then analysed according to the textual function of the antonymous pair. This analysis resulted in the development of the following eight distinct functional categories of antonymy:

1. Ancillary Antonymy
2. Coordinated Antonymy
3. Comparative Antonymy
4. Distinguished Antonymy
5. Transitional Antonymy
6. Negated Antonymy
7. Extreme Antonymy
8. Idiomatic Antonymy

Ancillary antonymy and coordinated antonymy account for 77% of the total number of sentences. In **ancillary antonymy**, the antonym pair functions as a lexical signal of a nearby contrast which is not necessarily entrenched in the lexemes used to express the contrast (e.g. *It is meeting **public** need not **private*** greed). This results in a sentence with two contrast relations in which the antonymous contrast underlines the contrast of what Jones calls the ‘B-Pair’.

**Coordinated antonymy** functions in a markedly different way. The antonymous pairs in these sentences indicate the inclusiveness or exhaustiveness of a scale and are usually conjoined by *and* or *or* (*the old as well as the young*, *both high and low incomes*). This is a relatively common phenomenon and some pairs behave this way exclusively (*confirm:deny; prove:disprove*) and, according to Jones, allow the contrastive power to stay dormant.

The other classes highlight comparison, contrast, change or movement, or negation of one member of the pair, and extremes of a semantic scale respectively. The last category of **idiomatic antonymy** includes phrases in which antonyms occur but which are essentially fixed idioms (*through thick and thin, easy come,* *easy go*). In my opinion these do not really constitute a class of textual functions as they are not constructed by combinatory means (on-line) but are always produced together and to a certain extent entrenched in the mental lexicon (or semantic memory) as fixed combinations.

For each of the above categories, Jones proposes certain sentential frameworks in which the antonym pairs commonly occur which can be used to search for certain antonym types in text and also to make tentative suggestions about potential future candidates for antonymy.

The classification aside, Jones investigates another feature of antonymy which is of interest to the questions raised in this study: the sequence in which antonyms occur. Antonyms seem to favour a particular sequence in text (and presumably speech) which leads to a marked and an unmarked order (*good:bad* vs. *bad:good*). Some antonym pairs included in Jones' study appear exclusively or almost exclusively in one order (*correct:incorrect; prove:disprove; rightly:wrongly; confirm:deny; agree:disagree*) whereas others proved more variable. A considerable number (17/56) indicate a statistically significant preference for a certain sequence of occurrence. This phenomenon of sequencing is a factor which has to be taken into account in experimental design as presenting an antonym in a marked order will presumably adversely influence measurements such as, for example, reaction time.

The criteria suggested for the phenomenon of sequencing are **morphology** (the derived antonym tends to follow the base one), **positivity** (the positive member of the pair tends to precede the negative one), **magnitude** (more precedes less), **chronology** (*begin:end),* **gender** (male before female), **phonology** (tendency to separate identical/similar syllables) and **idiomaticity** (frequent use in a certain combination). Frequency and markedness also seem to play a role, albeit one that cannot be statistically validated. The general tendency seems to be that the more frequent and the unmarked members of a pair precede the less frequent, marked ones. However, both of these criteria can be challenged by counter-examples: e.g. *old:new*; *new* is the higher frequency adjective (occurs three times as often as *old*) but in 72% of the cases of co-occurrence *old* precedes *new*. (Jones 2002: 130f.)

Jones cites *hot:cold* and *wet:dry* as problem pairs. However, in my opinion, if the magnitude criterion were rephrased to form a ‘more-or-less’ criterion, both those pairs would fit into that category, as *hot:cold* can be defined by the presence of more or less heat and *wet:dry* by the presence of more or less water/humidity.

Sequence is, according to Jones, also influenced by textual function. This can explain why the sequence is sometimes reversed. Syntactic distance seems to be an important criterion for sequencing. The closer together the antonyms appear in the text, the more likely they are to appear in their ‘usual’ sequence. In general, lexical antonyms are less likely to observe a certain sequence than morphological antonyms. The cases in which the sequence is reversed often give a certain rhetorical power to the sentence and draw the reader's attention to the antonym mentioned first (Jones 2002: 136).

Jones concludes his study with the question of whether word class and gradability affect the textual function of antonyms. He determines that gradability does not have an obvious effect on usage and that the similarities are more obvious than the differences.

The last part of the study is dedicated to the question of whether it is possible to determine which lexical pairs might become entrenched as antonyms in future. He proposes a search technique based on the frameworks drawn up for the categories of antonyms in his classification and then inserts word pairs to test their antonymy status.

The last question which is addressed in the final section is what antonymy truly is. It is proposed that both core vocabulary (Carter 1987) and semantic range (Muehleisen 1997) form part of the explanation, but the question of whether there is anything beyond co-occurrence and frequency which determines which pairs might be or become good antonyms is still unresolved. I will conclude this section in the same way Jones concludes his investigation: by giving his ‘new’ definition of antonymy which, slightly modified, will be my working definition for this current project (cf. Chapter 1):

> 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 (Jones 2002: 179).

## 2.3 The cognitive perspective

Croft and Cruse (2004: 164) start out by reminding the reader that ‘like all sense relations, oppositeness is a matter of construal, and is subject to cognitive, conventional and contextual constraints.’

Recently, with the increasing popularity of cognitive linguistics, sense/lexical relations have also begun to receive attention from this perspective. The most comprehensive account of antonymy in the cognitive semantic area has come from

Cruse and several colleagues of his, but has only treated antonymy (in the restricted sense) and complementarity.

Murphy (2003) has proposed a single relational principle in her discussion of lexical relations which she applies to all the relations which are said to structure the lexicon. She marks her investigation out as a pragmatic and psycholinguistic one. Her conclusions are of little relevance here but she does contribute to the canonical/peripheral discussion, which can be seen from the inclusion of a short discussion of her work below (2.3.2).

### 2.3.1 A cognitive model of antonymy

Cruse puts forward a suggestion for a model of antonymy based on the cognitive principles of schema (Lakoff 1987), domain and construal (Langacker 1987).[8] He explicitly states in Cruse & Togia (1995) that this is a development of previous ideas about antonymy put forward in Cruse 1986, 1992 and 1994 and extended further in Cruse 2000 and Croft & Cruse 2004. I will, in the following, mainly focus on Cruse & Togia's account (1995) as this is the most detailed.

Cruse and Togia (1995) only take into account the relation of antonymy in the restricted, Lyonsian sense, namely that of a pair of gradable, directionally opposed adjectives. Their main interest lies in the explanation of certain phenomena observed in the use of certain types of antonyms: committedness, impartiality and inherentness. All these are phenomena which have been of particular interest to scholars involved in the study of antonymy, as they are observable surface phenomena which lend themselves to classification. Even though it is not the area of my primary interest, this account, which is one of the few attempted in cognitive terms, illustrates the conceptual structure and reality of antonymy as a semantic relation rather than simply an association of certain lexemes.

As far as the cognitive concept of domain is concerned, Cruse & Togia differentiate between content domains, which bear a close resemblance to the world as it is, and schematic domains, which impose a specific construal on certain aspects of the content domain, but both are considered conceptual. Some types of construal fit better with some types of content domain than with others and therefore Cruse & Togia divide the schematic domain of antonymic construal into three major

[8]These three concepts are only explained insofar as they are directly relevant to the understanding of the theoretical approach proposed here. For a detailed account of all three see Ungerer & Schmid 2006.

types: monoscalar, biscalar and hybrid anto-complementary systems[9]. I will now introduce the first two systems and highlight how these relate back to the types of antonyms proposed by Cruse 1986. Furthermore, I will give a very brief account of how these construals achieve explanations for committedness phenomena observed in antonymy.

### (A) Monoscalar systems

Monoscalar systems are the conceptually simplest kind of construal and are usually associated with the class of polar antonyms in Cruse's 1986 classification (see 2.1.2). Therefore, the associated lexemes are objectively descriptive and the scale is usually calibrated in conventional terms. A simplified image of a monoscalar system is shown in Figure 2.2 below.

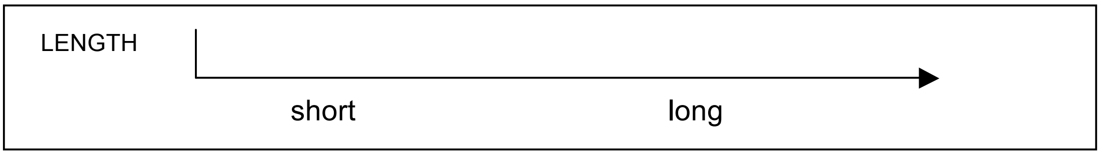



**Figure 2.2** *A monoscalar system*[10]

According to Croft & Cruse 2004, the association between certain scales and antonym pairs is strongly motivated. One reason for this assumption is strong cross-linguistic consensus as to which system of construal certain pairs of antonyms favour.

There are three reasons why a monoscalar construal is favoured in certain cases: salience of the properties, ease of construal of a determinate endpoint and the calibratability of the scaled property (Croft & Cruse 2004: 171).

If more of a property is seen as more salient, e.g. more length (*long*) is seen as more salient than less length (*short*), then this property is usually chosen as the basis of the scale (LENGTH instead of ?SHORTH). Furthermore there is, on a scale of LENGTH, an easily determined endpoint at the ‘short end’ but there is none at the other end; the scale can extend to an indefinite length. The third criterion refers to the calibratability of the scale: length is easily, and objectively, calibrated and generally, length is more easily calibrated than shortness. This shows an

9 These systems will not be discussed in detail here but include, for example, antonym pairs like *clean:dirty* and *tasty:bland*. For a detailed account see Cruse & Togia 1995.

10 Figures 2-5 are all modelled on Croft & Cruse 2004 (170f.).

overwhelming support for a scale of LENGTH.[11] The member of the antonym pair which represents the whole scale is usually the impartial term and also the term which, when intensified, denotes a higher value of the relevant property (*longer*), as opposed to its counterpart which will denote a lower value (*shorter*).

These scales can, in addition, be viewed and construed, in absolute terms (*15 km/h*, *3.5 kg*) or in relative terms (*a slow car*, *a heavy suitcase*) to a certain reference value of 'normal' cars or suitcases. These two interdependent scales account for the differences in use, and construal, of antonyms.

**(B) Biscalar systems**

Biscalar systems are defined as 'incorporating a pair of counterdirectional monoscalar systems' (Croft & Cruse 2004: 181). In Croft & Cruse 2004, they are divided into three distinct systems: disjunct equipollent systems (*hot:cold*), parallel equipollent systems (*hard:soft*) and overlapping systems (*good:bad*) whereas Cruse and Togia 1995 only differentiate between the overlapping and equipollent systems. The equipollent patterns, both parallel and disjunct, are completely symmetrical, either meeting at a zero point and extending into infinity in different directions, or in two parallel, equal scales running each other's full length. There are no system-impartial uses of either member of the antonym pair in equipollent systems. This means that both terms are supra terms where the comparative denotes a greater degree of the property expressed (e.g. *colder*: more coldness vs. *hotter*: more hotness) as opposed to the existence of a supra and a sub term. The overlapping system consists of two scales which are not equal, with one running part of the length of the other. Here, only one of the two members has system-impartial application.

The question which arises with the equipollent systems is why there are two scales involved and why they are positioned end-to-end. Cruse & Togia give the following, very accessible, explanation:

> Our answer is that that is the most comfortable way for scale schemas to fit onto the content domain. Imagine putting one's hand into (i) a bowl of cold water, (ii) a bowl of tepid water and (iii) a bowl of hot water (leaving time for one's receptors to go back to their resting state in between each). One will get a strong temperature sensation in (i), no temperature sensation at all in (ii), and a different strong temperature sensation in (iii). In other words, there is a basic natural zero sensation in the middle, and two distinct salient temperature sensations: a 'natural' equipollent situation. (1995: 116).

[11] SPEED and DIFFICULTY are other examples cited in Cruse & Togia 1995 and Croft & Cruse 2004 which are marginally less clear (especially in the case of DIFFICULTY) but still convincing examples of monoscalar systems.

However, *hot:cold* still operate in the same system, as we all know from experience that if we add hot water to cold it will gradually become hotter. Therefore, a construal as depicted in Figure 2.3 seems appropriate even though this may appear somewhat counterintuitive as the scales are not connected whereas the temperature continuum is perceived as a whole.

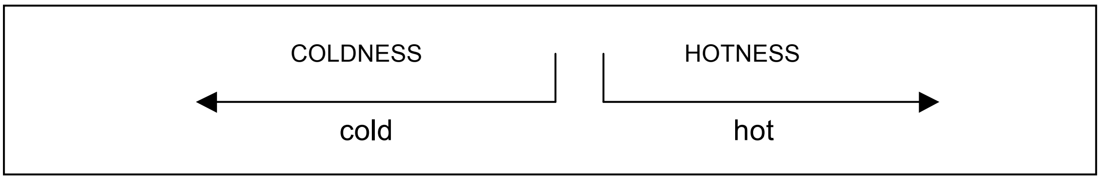



**Figure 2.3** *A disjunct equipollent system*

There are not many disjunct equipollent pairs in English, and initial investigations suggest that they might not exist in other languages (e.g. Turkish) at all. Examples in English are: *happy:sad, sweet:sour, ashamed of:proud of* and *beneficial:harmful* (Croft & Cruse 2004: 182). Parallel equipollent pairs (e.g. *hard:soft, dark:light*) are even rarer and they occur when '(a) there is no property that can be construed as having a zero value in the middle of the scale, (b) each direction of construal is equally motivated and (c) there is no dominant viewpoint, no difference that motivates a 'positive/negative' construal' (Croft & Cruse 2004: 182f.).

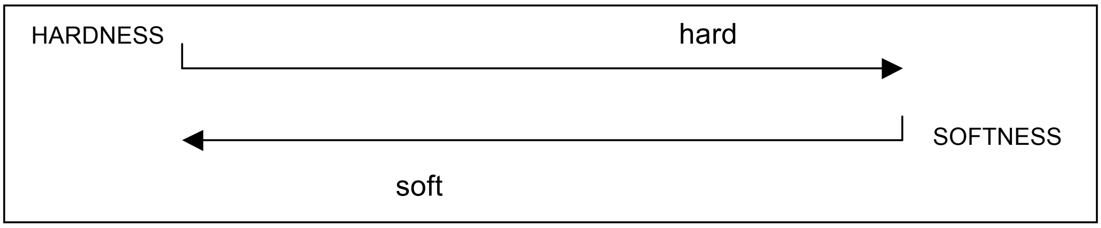



**Figure 2.4** *A parallel equipollent system*

Overlapping systems, where a second scale runs along part of the primary scale, are fairly common and are usually construed for the category of overlapping antonyms in Cruse 1986. Examples of these are *cruel:kind* and *polite:rude*: pairs which usually encode some kind of positive/negative distinction aligned with a socially accepted and conventional norm, rather than an objective, calibrated quantity as in the case of monoscalar systems.

Inherentness plays a role in the construal of overlapping antonyms as some lexemes have inherent qualities which prohibit collocation with e.g. *better* in certain situations (cf. Cruse & Togia 1995: 135). An interesting idea is that of a certain type of implicature of placing something on the smallest possible scale which encompasses all the possible outcomes of the reply: a question like *How good was the film?* indicates that the speaker is prepared for any answer on the MERIT scale. However, if we ask *How bad was the drought?* only replies on the BADNESS scale are considered, as a drought is inherently a bad thing. In my opinion, however, the question *How bad was the film?,* which does not have the intrinsic quality of badness adhering to the lexeme film, similarly restricts the answer to the BADNESS scale, as the speaker expects a reply which will not extend into the positive part of the MERIT scale.

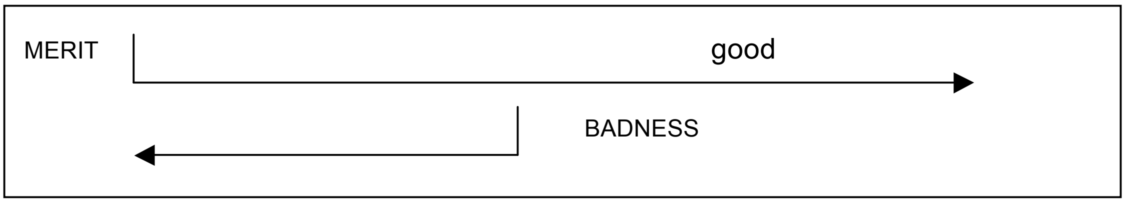



**Figure 2.5** *An overlapping system*

One last interesting hypothesis put forward by Cruse & Togia (1995) and Croft & Cruse (2004) is the **exposure hypothesis**. They state that 'there are certain notions that have a high priority for expressibility in languages and which will tend to be assigned as the default readings of certain expressions' (Cruse & Togia 1995: 127). In these notions they include impartial expressions of calibrated degree properties, and relative quantities and impartial comparisons of quantities. The exposure hypothesis suggests that 'the more exposed an adjective the greater its affinity for the relative scale, and the less exposed, the greater its affinity for the absolute scale' (Cruse & Togia 1995: 128). However, they issue the caveat that this is only a tendency.

They suggest the following degrees of exposure:

(i) cases of suppletion: *speed, weight, temperature, worse, better*
(ii) cases of morphological distortion: *length, width, depth*
(iii) undistorted incorporation in a derived word: *thickness, hardness, difficulty, longer, thicker*
(iv) expressions where the bare adjective occurs in an idiomatic construction: *How long is it?*
(v) bare adjectives in fully compositional expressions: *It's long, a long piece of string*

(taken from Croft & Cruse 2004: 177)

The following examples show varying degrees of exposure and therefore varying degrees of committedness and preference for the absolute or relative scale respectively.

(2.8a) *I was surprised at its* ***quickness****.*
(2.8b) *I was surprised at its* ***speed****.*
(2.8c) *I was surprised how* ***slow/fast*** *it was.*

(2.8b) is impartial and favours a construal on the absolute scale, whereas (2.8a) is committed but could be construed on either a relative or an absolute scale. (2.8c), however, is completely committed and can only be construed on a relative scale.

These investigations show that cognitive constraints have an influence on the properties antonyms display on the surface and which might seem, at first glance, rather idiosyncratic. The importance of cognitive criteria should not be underestimated but it remains to be seen whether these findings can also be applied to more peripheral phenomena of opposition with equally informative results as these phenomena might be, at least cognitively, less rigidly structured. However, as Cruse & Togia themselves suggest, putting their theory to the test in ‘an investigation of the contextual sensitivity of antonym behaviour, especially those outside the core group of polar antonyms which form the focus of most investigations of antonymy’ (Cruse & Togia 1995: 141) seems a worthwhile objective from a cognitive angle.

### 2.3.2 A cognitive-pragmatic approach (Murphy 2003)

Murphy has published several articles on antonymy (1993, 1994, 1995 and 2000), preceding her publication on semantic relations and their representation in general. Consequently, she places most emphasis on antonymy and large parts of the general discussion are based on her previous research on opposition.

Her overarching goal is ‘to provide an account of how individuals know (or determine) whether words are semantically related or not and, if they are related, what type of relation is involved’ (Murphy 2003: 4). She subscribes to a psycholinguistic-pragmatic approach and divides her overall goal accordingly: on the psycholinguistic side she aims to ‘provide a psychologically plausible model of the knowledge or processes involved in semantic relations phenomena in human language behaviour’ (Murphy 2003: 4f.). The pragmatic side is accounted for by the fact that the phenomena being discussed are ‘considered in reference to their use and their status in a human mind with human culture’ (Murphy 2003: 5).

As far as the pragmatic angle is concerned, Murphy's explanations and conclusions are rather unconvincing as they focus mainly on the context-dependence of lexical relations which, in her account, seems to be highly idiosyncratic. I fully support a contextual explanation of certain phenomena within antonymy (further work on contextual aspects of opposition for example in Murphy & Andrew 1993, Charles 2000) but strongly disagree with Murphy's approach of substituting a highly idiosyncratic approach for the 'ordered' flexibility which a contextual application allows, and in this study instead support the idea that there are certain pragmatic principles at work which are 'rule-governed' rather than random, as this can account for not storing every separate instance of an antonym pair in our mental lexicon.

Murphy further claims that lexical relations cannot be innate as they are language specific, a point few scholars would disagree with, and joins the chorus of semanticists deliberating whether semantic relations are indeed semantic and generated by a set of rules or whether they are an associative relation and are learned by co-occurrence.

Canonical antonyms are defined as ‘the kinds of antonyms that automatically follow one another in free word association tasks, or that are collected together in children's books of opposites’ (Murphy 2003: 10) and non-canonical antonymy is characterised, according to Murphy, by lower frequency or greater context dependence. These concepts are discussed further, but as an initial assessment of the concepts this is a rather unorthodox definition. Children's books or games for antonyms do indeed gather mainly canonical antonyms but there are also ones which might not be considered antonyms at all, e.g. *loose:fit*, *sweet:bitter* or *few:more*. Murphy does, however, argue for a theory of semantic relations which takes into account the judgements made by speakers about ‘better’ and ‘worse’ examples of a semantic relation (cf. Murphy 2003: 11f.).

An interesting point made in connection with canonicity is the idea that antonym pairs which are stable across various readings of the lexemes involved (e.g. *hot:cold*) display a greater degree of canonicity than others. This would bolster the co-occurrence hypothesis to some degree, as antonyms which are opposed in more than one meaning will co-occur even more frequently because they occur together in several meaning-combinations, rather than just one. This notion of stability is also one which is incorporated in the experimental design of the questionnaire discussed in Chapter 3.

The core of Murphy’s study is the introduction of a single relational principle which determines all semantic relations:

**Relation by contrast (RC)**

The contrast relation holds among the members of a set iff: they have all the same contextually relevant properties but one (Murphy 2003: 44).

This principle is a relational principle which functions on the basis of minimal difference and states that members among relational sets of all kinds (e.g.

synonymy, antonymy, meronymy etc.) are minimally different, albeit in different ways. For antonymy, this relational principle is extended into RC-LC (relation by contrast – lexical contrast) which is seen as a relation between 'word-concepts'. Murphy dismisses the finely differentiated categories drawn up by Cruse (1986) as irrelevant and deems that nothing is gained from an approach like this as most of the properties introduced in Cruse's account can be subsumed under RC-LC. Instead she advocates a general definition which would allow for all opposites and the prototype effects displayed by this category (cf. 6.1.3).

Murphy also discusses markedness effects and the influence of binarity, both of which are central concepts to the research undertaken here; these will be discussed later in the appropriate context (cf. Chapter 6).

While Murphy's study takes into account many of the phenomena I am interested in – context-dependence, cultural variability, base antonyms and many others – she does not use experimental data to substantiate her claims and theories, and part of what I will undertake in the present study follows on from ideas and claims put forward in Murphy 2003.

## 2.4 The psycholinguistic perspective

Antonymy, and sense relations in general, have also been a focus of interest of psycholinguists, psychologists and neurolinguists as any indication as to how words might be stored in the mental lexicon is of great interest. Elaborate systems have therefore been developed to discover more about the structure of how words are stored in the mind. I will, in the following sections, look at one approach in particular, the WordNet model, and then go on to outline previous psycholinguistic research done in the area on which my study will partly be based as this will aid in providing the context of my own research as well as outlining what has already been achieved.

### 2.4.1 A psycholexicological approach – WordNet[12]

WordNet was designed by a research group at Princeton University with the aim of organising the English lexicon into a system which is guided by current knowledge

[12] This section is based on Miller 1990 and Gross & Miller 1990 and reflects the state of WordNet at the time of writing.

and assumptions about the way lexical and semantic information is stored in our minds. It is therefore less cumbersome and more intuitive to use than the traditional alphabetical dictionary (cf. Miller 1990: 235f.) This, of course, has only become possible thanks to fairly recent advances in technology.

I will briefly outline the general structure of this lexical reference system before looking specifically at the treatment of adjectives in WordNet, as this is the most relevant aspect to antonymy, which is assumed to function as a structuring principle of the adjectival lexicon.

WordNet contained (in 1990) roughly 54,000 distinct lexical entries arranged in synonymy sets (synsets); many psycholinguistic hypotheses about the structure of the mental lexicon had to be discarded as a larger sample of language was considered. WordNet divides the lexicon into four categories: nouns, verbs, modifiers and function words, but only contains nouns, verbs and adjectives. Function words are omitted due to the common assumption that they are probably stored separately, as they seem more closely related to syntax. Adverbs are also not included as many of them are derived from adjectives and the creators of WordNet were intent on minimising redundancy.

There seems to be 'no single organising principle in the lexicon' (Miller 1990: 237) as nouns seem to be organised by topical hierarchies, verbs by entailment relations and adjectives (possibly) by opposition. However, WordNet is not interested in organising lexemes by word form but rather by word meaning, and the multitude of mappings between forms and meanings presents a problem for both speech production and speech processing. Synonymy and polysemy are seen as two sides of a coin: while synonymy is considered a problem in the production of language, polysemy is taken to be a problem for language processing and comprehension.

As mentioned above, WordNet is organised in synsets and semantic relations are considered pointers between synsets. The organisation of the synsets is used to automatically disambiguate polysemous lexical items as these will appear in both relevant synsets. The relations used in the WordNet framework are synonymy, antonymy, hyponymy and meronymy. The remainder of the discussion here will focus on the organisation of the adjectival lexicon as the relation which is considered the pointer between adjectival concepts (or lexical items) is antonymy.

The creators of WordNet take the stance that 'antonymy is a lexical relation between word forms, not a semantic relation between word meanings' (Miller 1990: 242). It is nevertheless considered the central organising principle of the adjectival lexicon in WordNet, despite the creators' explicit focus on meaning relations. The class of adjectives is divided into **ascriptive**, 'one that ascribes a value of an attribute to a noun' (Gross & Miller 1990: 266) and **non-ascriptive** adjectives (which the creators of WordNet have affectionately named '**pertainyms**' (from 'pertaining to')), for example, *a musical instrument* which is not *an instrument that is musical* but *an instrument used in music* (Gross & Miller 1990: 269). There are, of course, adjectives which can be found in both ascriptive and non-ascriptive functions (e.g. *a nervous person* vs. *a nervous disorder*). I will focus on ascriptive adjectives as these are the group which appears to be structured by antonymy.

Figure 2.6 below shows a model of how adjectives could be related in WordNet. The principal relation here is direct antonymy (between *wet:dry*; symbolised by the arrow with the broken line) and synonymy functions as a secondary relation (arrows with unbroken line) linking those adjectives without antonyms to ones which have a corresponding antonym to thus link opposing concepts. This system relies on a clear distinction between semantic and lexical relations. The principal connecting devices are familiar (canonical) antonym pairs (lexical relation) whereas the conceptual relations (semantic) are often shown as being mediated by the more salient lexical relation. This model suggests (and experimental work has been carried out in this area) that indirect antonyms, for example *soggy:parched*, are processed via the direct antonym pair *wet:dry,* thus explaining the longer processing time required by indirect relations. The experimental studies on indirect antonyms and their results are discussed in 2.4.2.2.

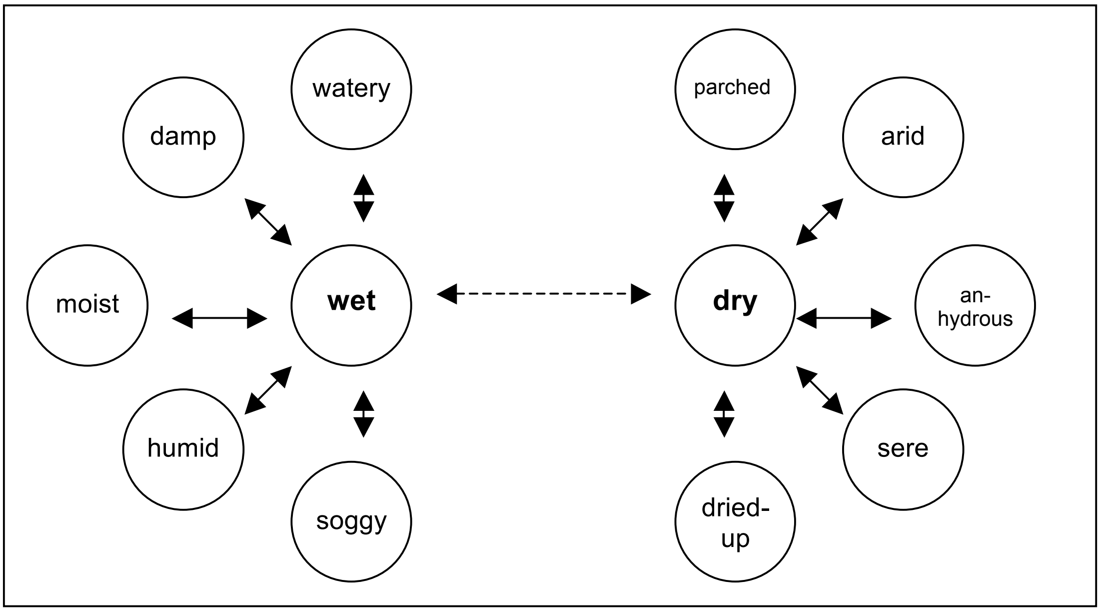



**Figure 2.6** *Bipolar adjective structure in WordNet (Gross & Miller 1990: 268)*

There is, however, another antonym pair in the figure above which seems to warrant the designation of direct antonym: *humid:arid*. These multiple direct antonyms within the same synset (also *little:big, large:small*) cause problems for the construction of synsets as there are two candidates for the primary relation which are, in some cases, equally eligible (cf. 3.4.1.3 & 4.4.2.2). The solution, in case of two equally salient pairs, is to create overlapping clusters (Gross & Miller 1990: 276).

The creators of WordNet have taken into account phenomena such as gradation and markedness but have not coded them in WordNet. The only criterion coded in WordNet which is pertinent to my research is that of selectional preferences. This refers to the preferences of collocation and the question of how it is determined which adjectives will be able to modify certain nouns. Gross & Miller (1990: 274) ‘assume that the interactions between adjectives and nouns are not prestored, but are computed as needed by some on-line interpretative process’. Thus, nouns and adjectives are coded in certain ways to determine their compatibility.

The WordNet model has given rise to much discussion and, as well as being based on experiments, has raised additional questions which could be answered by experimental means.

### 2.4.2 Experimental approaches

Lexical relations in general and opposition in particular have been experimentally investigated from various angles (among others Charles & Miller 1989, Gazzaniga & Miller 1989, Gross et al. 1989, Murphy & Andrew 1993, Charles et al. 1994)[13] and many of these studies have given rise to considerable debate about the nature of the relation involved in antonymy. This, however, will be discussed in 2.5.3 in the discussion on frequency of co-occurrence. In this section, I will briefly sketch early word association experiments relevant to antonymy and then, in 2.4.2.2, introduce some more recent experimental studies.

#### *2.4.2.1 Laying the foundations – early word association tasks*

Word association experiments in relation to antonymy have been carried out in various psycholinguistic contexts with varying perspectives and results. Deese (1964/1965) was one of the first to study antonymy in a systematic way within his study of the organisation of word meanings in the mind. He used free association to

[13] A number of more recent studies which focus more on the neurological or psychological mechanisms underlying the processing of lexical relations are introduced in the context of the behavioural experiments in Chapter 4.

elicit a 'related' word without putting any constraints on the responses the informants were allowed to give. Deese discovered, among many other things, that the common adjectival antonyms are often very strongly associated with each other and this led him to put forward the idea that the adjectival lexicon of English might be structured by antonymic relations.

> The strong notion of contrast implies that one member of a pair of words should have its associative meaning most strongly determined by the other member of the pair and that the meaning should be reciprocal. Contrasting words, in brief, are words that are defined in terms of each other [...]. (Deese 1965: 122)

Deese went on to determine the best antonym pairs by testing their reciprocal associative strength and stated that in each case in which one word is the most frequent response to the other and vice versa, that pair can be considered a basic antonym pair. His list of antonyms, which I have also taken into consideration in designing experimental tasks, is reproduced in Table 2.1.

**Table 2.1** *Deese's 39 basic antonym pairs*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| alone | together | dark | light | hard | soft | old* | young |
| active | passive | deep | shallow | heavy | light* | poor | rich |
| alive | dead | dry | wet | high | low | pretty | ugly |
| back | front | easy | hard* | inside | outside | right | wrong |
| bad | good | empty | full | large | small | rough | smooth |
| big | little | far | near | left | right* | short | tall |
| black | white | fast | slow | long | short* | sour | sweet |
| bottom | top | few | many | married | single | strong | weak |
| clean | dirty | first | last | narrow | wide | thick | thin |
| cold | hot | happy | sad | new | old | | |

According to Deese (1965), the lexemes marked with an asterisk (*hard, light, old, right, short*) did not display the same level of associative strength displayed by the other pairs as each of them has more than one common antonym which could be elicited and which therefore lowered associative strength results for the competing lexemes. In all instances, however, both pairs displayed enough associative strength to be included in the list.

The list in Table 2.1 still serves as the basis for many experiments and studies (e.g. Jones 2002) when a group of antonyms needs to be considered. The results obtained by Deese laid the foundation for research into canonical and non-canonical antonyms and highlighted this special relation between central members of the category of antonyms which has been debated extensively. Furthermore, Deese drew attention to the fact that two words become associated by sharing context but did not pursue the matter further (cf. Deese 1965: 128).

*2.4.2.2 Conceptual vs. associative processing – more recent evidence*

Charles, Reed and Derryberry (1994) carried out two sets of experiments which were designed to determine whether antonymy was processed in an associative or conceptual manner and whether there was a processing difference in indirect vs. direct antonyms. The debate whether antonymy is a conceptual or an associative relation determined by frequency of co-occurrence is also discussed further in Chapter 3.3.3 below (as well as in Chapters 4 and 5 in relation to more experimental data and in Chapter 6 from a theoretical perspective).

Charles et al. (1994) started out from the evidence gained from Deese's experiments as well as hypotheses of co-occurrence put forward by the creators of WordNet (e.g. that the adjectival space is structured by antonymy which is, in this case, seen as an associative relation between certain lexical items rather than between opposed concepts). One condition all views seem to agree on is that direct, or canonical, antonyms have some special property which connects them very closely and accounts for their high associative strength in word association tasks (cf. Deese 1964/1965, Jenkins & Palermo 1964), their positive priming effects (Becker 1980) and their speedy identification as antonym pairs in reaction time experiments (Hermann & Chaffin 1986, Gross et al. 1989, Charles et al. 1994).

There are, however, varying strengths of this argument. Gross et al. (1989) and Charles & Miller (1989) put forward the argument in its purest form, namely that direct antonyms are learned through frequent co-occurrence rather than associated through strong conceptual links and thus more fundamental to the organisation of the adjectival lexicon than semantic similarity. Gross et al. (1989) explained the faster decision times for direct antonymy regarding their relatedness by stating that the conceptual opposition task is slower where indirect antonymy is concerned as the decision must be mediated by the direct antonym pair (cf. Charles et al. 1994: 330f.). This also explained those of their results which showed that the more distant the indirect antonym pair was from the related direct antonym pair, the longer the decision time overall.

Another study suggests a somewhat more cognitive approach to this question: Gazzaniga & Miller (1989) carried out a neuropsychological study of a split-brain patient in which they tested the patient's ability to recognise direct and indirect antonyms. Here, the advantage in processing for direct antonyms is ascribed to a link between the direct antonym pair and the related attribute dimension, as the patient did not use lexical knowledge to describe attributes. The patient's right hemisphere showed no difference in reaction to direct and indirect antonyms whereas the left hemisphere responded normally. The results of the experiments pose problems for the co-occurrence hypothesis as there is no difference in the processing of the pairs in the right hemisphere of the patient which would have been predicted by supporters of the co-occurrence hypothesis since they propose different types of processing for direct and indirect antonymy. Furthermore, it highlights the problems with the approach put forward by the WordNet creators that indirect antonymy is mediated by direct antonymy as the right-hemisphere results show that it is not impossible to understand indirect antonyms without knowing the direct antonyms which are supposed to function as a link (cf. Gazzaniga & Miller 1989: 192).

The third view in this discussion, represented, for example, by Murphy & Andrews 1993 (cf. also 3.3.3), is a purely conceptual one. They argue that antonymy is a relation between opposed concepts and that direct antonyms do not need to be considered ‘special’ as far as their type of relation is concerned. They include stylistic, connotational and morphological factors in their explanation. Their experimental research required informants to give antonym responses either in isolation or in a noun phrase and measured the degree of agreement between the two contexts. When provided with the lexeme *fresh* in three contexts: *fresh shirt*, *fresh idea* and *fresh fish* (another possibility would have been *fresh bread*) the antonyms provided did not match. The replies given were, for example, *dirty shirt*, *old idea* and *frozen fish* (presumably a fourth reply would have been *stale bread*) (cf. Murphy & Andrew 1993: 306). This shows that antonym pairs, even canonical ones, are strongly context-dependent (cf. Chapter 6).

The assumption was that if the antonym reaction was indeed determined by an associative link, the contextual effects would be minimal. However, the responses given varied greatly with regard to context, and the overall agreement between isolated pairs and pairs in noun phrases was fairly low.

Based on these three standpoints, Charles et al. designed their experiments[14] on the hypothesis that ‘if antonymy is essentially based on lexical associations and synonymy is essentially conceptual, it should be possible to find processing differences between the two types of relationships’ (Charles et al. 1994: 332). The authors worked with a paradigm that measured the difference in speed

[14] Much of the experimental design and the details of the methodology will be presented at a later stage when the methodology used in the present work is introduced (cf. 4.2).

with which informants react to three different kinds of relations: direct antonyms (e.g. *hot:cold*), indirect antonyms (e.g. *hot:frigid*) and synonyms (e.g. *hot:warm*). An adjective prime was presented and then followed by a second adjective 350ms later. The informants were, in this instance, only required to decide whether the adjectives presented to them were related or not: they pressed one key for ‘related’ and another for ‘unrelated’. Lexical markedness of the prime and semantic distance was counterbalanced among the informants.

Related direct antonym pairs (both marked and unmarked) elicited both the fastest response time among all stimuli and the most accurate responses. This supports an associative model of antonymy. However, synonyms and indirect antonyms showed similar semantic distance effects, which is consistent with the conceptual approach to indirect antonymy.

The second experiment carried out by Charles et al. (1994) was designed to test the processing speed of direct synonyms and direct antonyms. Here, the informants were presented with two tasks: a synonym task in which they had to decide whether the given pair consisted of two synonyms or two unrelated words and an antonym task following an analogous procedure.

According to Charles et al. (1994: 346), the findings of both experiments suggest three important effects: ‘(1) highly efficient processing of direct antonyms; (2) more efficient processing of synonyms than indirect antonyms; and (3) similar semantic distance effects for synonyms and indirect antonyms.’ The results also support the theory that associative links are involved in the processing of direct antonyms, but they do not suggest that it is only the associative link which creates this ‘special’ relation.

They conclude with the following remarks:

> The first interpretation is that conceptual processes and attentional mechanisms may account for these data. Semantic similarity provides a good explanation for the close and far antonym differences, as well as for speed differences between direct, close and far synonyms. […] A second interpretation takes an associative position. […] A third interpretation incorporates elements from both of these interpretations. Direct antonyms appear to be different from both synonymy and indirect antonyms in their highly efficient processing. (Charles et al. 1994: 350).

I will, in this study, start out from the third statement, as it is impossible to disregard the effects of entrenchment which frequent co-occurrence, especially in a limited number of syntactical patterns (cf. Mettinger 1994, Fellbaum 1995, Jones 2002), has on associative strength. However, on the basis of the close analysis of several cases of peripheral antonymy, I will also argue for a strong conceptual link which contributes significantly to the relation of direct antonymy.

## 2.5 Direct antonymy – what makes an antonym pair canonical?

All of the accounts discussed above divide opposition into good and bad opposition; in other words into ‘canonical’ and ‘peripheral’ or ‘systemic’ and ‘non-systemic’ antonymy. This is in some accounts largely based on individual native speaker intuition and in others on the results of psycholinguistic experiments such as Deese's word association experiments (1964) or the reaction time experiments on direct and indirect antonyms (Charles et al. 1994; cf. 2.4.2.3)

The question of what makes a ‘good’ antonym is, however, only rarely discussed explicitly. Cruse (1986: 262), for example, spends only one paragraph on this particular question at the end of his three chapters on opposition, and many others assume this distinction to be common knowledge which requires no specific justification.

This section will attempt to draw together features which have been put forward as characterising canonical antonymy and will illustrate these using the example of the temperature scale including the following lexemes:

| |
|---|
| boiling – **hot** – warm – lukewarm/tepid[15] – cool – **cold** – freezing |

**Figure 2.7** *The TEMPERATURE scale*

Any pairs of the lexemes above could technically be considered opposites as they form three pairs which all stand in a certain opposition to each other. However, there is initially no reason why there should not be other pairs; *hot:tepid*, *lukewarm:cold*, *freezing:warm*, *tepid:freezing* and *boiling:warm,* to name but a few examples. All of these incorporate a certain degree of opposition as they cover different sections of the temperature scale. However, every native speaker of English will have specific intuitions which will lead them to the conclusion that some of these lexemes pair up with certain other ones.

Of the lexemes mentioned in Figure 2.7 above only *hot:cold* are considered canonical antonyms while *lukewarm:tepid* is generally not even judged to be an antonymic relation at all. However, for illustrative purposes, I will assume for the moment that it exemplifies the antonymic relation which displays the least difference in heat on the temperature scale.

[15] *Lukewarm:tepid* is treated as a potential symmetrical antonym pair in parts of the discussion here despite the fact that, strictly speaking, both lexemes encode WARMTH and can mostly be used interchangeably.

I will now examine each of the criteria put forward in turn and attempt to illustrate how immediately and intuitively useful each seems to be. Features like **binarity** and **gradability**, which have already been addressed in the discussion of classifications of antonymy in 3.1, will not be discussed again here.[16] The first two sections will deal with semantic properties of antonym pairs whereas section 2.5.3 will take a closer look at the frequency of co-occurrence question and the issue of whether antonymy should be considered a conceptual or an associative (often also semantic or lexical) relation.

### 2.5.1 Cruse's criteria (Cruse 1986)

First of all, I will introduce the three criteria Cruse proposed in his brief discussion of what makes a ‘good’ antonym. He does not restrict these characteristics to antonymy in the narrow sense but simply postulates three criteria which, in his opinion, contribute to making an opposition a ‘good’ opposition (Cruse 1986: 262). Before listing the three factors which he considers worth discussing he states that ‘a binary directional opposition must form part of the meaning of any pair of opposites [and that] … part of their meaning must be at least to some degree patent.’ (Cruse 1986: 262). This refers back to Cruse's criterion of inherent binarity mentioned above (2.1). He notes that good opposites must be not just binary but inherently binary, as almost all pairs of co-hyponyms can be construed in a binary way in certain contexts.

#### *2.5.1.1 The uni-dimensional scale*

‘The first [factor] is the ease with which a uni-dimensional scale can be conceptualised, on which the opposed terms may be symmetrically disposed’ (Cruse 1986: 262).

There are, in my opinion, two essential factors contained in the above quotation: firstly, it is essential that a **uni-dimensional scale** can be established relatively easily for any given opposition as the failure to construct such a scale would lead to the opposition being judged less good. The second factor concerns the distribution of the opposite members on the scale; according to Cruse (1986, Cruse & Togia 1995), and many others, this distribution should be **symmetrical** and equidistant from the mid-point of the scale (cf. also Lehrer & Lehrer 1982).

[16] For a detailed discussion of binarity and the importance of gradability in antonymy see Cruse 1986 (Chapters 9 & 10) and Lyons 1977.

There are many examples which spring to mind to which these criteria are easily applied: *big:little* which operates on a scale of size, *wet:dry* on a scale of wetness, *fast:slow* on the scale of speed, and the example of the temperature scale introduced above, to mention but a few.

Considering the example of temperature, it is easy to see that a uni-dimensional scale is constructed without difficulty and that, if the symmetrical distribution is taken into account, four antonym pairs can be constructed (*boiling:freezing*, *hot:cold*, *warm:cool* and *lukewarm:tepid*) which are all equidistant from the midpoint of the scale.

Cases in which it is difficult to extrapolate one scale are, for example, *work:play* and *town:country*, examples which are not usually considered opposites unless in a very specific context as, according to Cruse, there are too many variables involved in the construal. Therefore the criterion of minimal difference, which is inextricably linked to the construal of a uni-dimensional scale, does not hold. These examples nevertheless contain a certain amount of opposition, and in certain contexts a case can be made for them to be counted as antonyms. However, they also fail to meet the criterion of inherent binarity.[17] This criterion alone, however, is not able to differentiate between the different degrees of 'goodness of opposition' perceived in the four pairs above. In a goodness-of-exemplar rating task, I assume, the ranking would be as follows: *hot:cold*, *warm:cool*, *boiling:freezing* and finally *lukewarm:tepid*. Nevertheless, according to the criteria mentioned so far, all the above are still considered equally good pairs of antonyms.

#### *2.5.1.2 Purity of opposition*

The second factor Cruse introduces is **purity of opposition**. He claims that there are opposites which are purer and therefore presumably more basic than others:

> [….] what proportion of the meanings of the opposed terms is exhausted by the underlying opposition: the greater this proportion is, the stronger the felt opposition will be. This is why *father* and *mother* are weaker opposites than *man* and *woman*, which in turn, are weaker than *male* and *female* […] (Cruse 1986: 262)

Greater complexity, and therefore more information that requires processing, is here seen as detrimental to the recognition and judgement of the antonymic relation concerned. *Male* and *female* are so-called 'pure' opposites as there are no additional facets of meaning to be processed apart from the basic dichotomy.

---

17 The context-dependence of antonym pairs and the construction of relevant contexts will be discussed in detail in Chapter 6.1.3.

**Table 2.2** *Feature semantic analysis of gender lexemes*

| | [+/- MALE] | [+/- HUMAN] | [+/- ADULT] | [+/- PARENT] |
|---|---|---|---|---|
| male | + | | | |
| female | - | | | |
| man | + | + | + | |
| woman | - | + | + | |
| father | + | + | + | + |
| mother | - | + | + | + |

The concept of purity of opposition is possibly one which is most easily expressed using semantic feature theory (cf. Table 2.2). Although all six lexemes, and consequently all three pairs, contain feature [+/- MALE] as the only distinctive one, the pairs which possess the smallest number of additional features are the purest relations.

Returning to the example of the TEMPERATURE scale, the concept of purity of opposition does not apply as all four pairs display a relatively pure opposition, only differing in the degree of temperature. However, there are cases corresponding to the examples above which apply to the temperature scale: consider, for example, the opposite pairs *fire:ice* or *summer:winter*. In both cases the distinction between *hot* and *cold* is one of the central ones involved, but there are other factors which obscure the temperature opposition. *Summer:winter* is, admittedly, not an ideal example, as it is a very good antonym pair due to its additional antipodal opposition. The *hot:cold* relation plays, in this case, a less important role.

Muehleisen (1997: 14) adds, in her discussion of Cruse's concept of purity of opposition, that, as what Cruse considers pure antonyms are exclusively adjectives and the impure opposites are nouns, there might be a relationship between word class and the purity of opposition which remains to be investigated. It seems logical that nouns usually carry more features than adjectives and are therefore more complex. However, there are also adjectives which are more complex than those considered so far and cannot be reduced to a simple distinction along one axis (cf. Chapter 3).

#### *2.5.1.3 Non-propositional meaning*

The third, and last, factor Cruse mentions is a **close match in non-propositional meaning**. Other than the following brief explanation, this point is not further illustrated but it is picked up in a significant number of publications (most notably Murphy 1990, Lyons 1995, Muehleisen 1997 and van Jaarsveld & Draskovic 2003).

> [A] good pair of opposites must be closely matched in respect of their nonpropositional meaning: that is why, for instance, *tubby* and *emaciated* are not fully satisfactory opposites, although they incorporate a binary directional opposition. (Cruse 1986: 262)

Here, it will simply be taken to mean that the two lexemes concerned have a similar range of application and similar collocational restrictions. *Tubby* and *emaciated* are, therefore, not a good example of opposition as they are used in clearly distinct registers and also have extremely different connotations. Another factor which comes into play here is collocational range, which is to say the number of lexemes a member of an antonym pair, or indeed any lexeme, co-occurs with. The more broadly applicable both members of an antonym pair are, the more 'general', the better an example of opposition they usually are. There are exceptions but these are usually not very common lexemes and therefore not known to a large number of informants, which makes them extremely difficult to test. One such example would be the antonym pair *recto:verso* which is used to refer to the right or left side of a publication but is hardly ever used[18].

Considering the temperature scale once again, it is immediately evident that the pair with the greatest collocational range is *hot:cold*, followed fairly closely by *warm:cool.*[19] *Boiling:freezing*, apart from its literal sense which is restricted to liquid substances (*boiling/freezing water*), is used with air temperature (*a freezing room*, *It's boiling today*) or body temperature (*freezing hands*) and *tepid:lukewarm* only refers to liquid substances (*lukewarm tea*, *a tepid bath*).

The pairs which are deemed less good examples of opposition, according to this criterion, seem to be consistent with the earlier intuition that *hot:cold* and *warm:cool* are somewhat better examples of opposition than the other two. There are, however, still other criteria to come which were not considered by Cruse or which, like his third criterion of non-propositional meaning, have been greatly extended by other scholars. These will be discussed in the following sections.

### 2.5.2 Semantic range (Muehleisen 1997)

Muehleisen's concept of semantic range corresponds to a certain extent to Cruse's non-propositional meaning discussed above and also to Lehrer & Lehrer's (1982)

[18] Originally, *recto* and *verso* are the terms used for the front and the back of a piece of parchment.
[19] For a more detailed analysis of the adjectives on the temperature scale see 3.4.1.1 and 5.2.2.1.

concept of distribution and Lyons' (1995) application of collocational restriction to synonymy and antonymy. A substantial number of studies have looked at the combination of adjective-noun phrases and the insights gained from these into lexical relations, especially synonymy and antonymy.

My reasons for using Muehleisen's term 'semantic range' rather than any of the other concepts are that it encompasses the other concepts proposed and it is the most recent approach, which therefore draws on the most up to date literature in the field. Furthermore, her study of two antonym fields (*big:little/large:small*, *wet:dry* and *happy:unhappy/sad*) is one of the most detailed investigations into the concept of semantic range.

Muehleisen begins her study with the assumption that ‘good opposites also describe the same kinds of things’ (Muehleisen 1997: 113). Her data is collected from several sources: learners' dictionaries (such as the Oxford Advanced Learners' Dictionary (OALD), the Longman Dictionary of Contemporary English (LDOCE) and the Longman Language Activator (LLA)) but also from a corpus only tagged for part of speech (POS) consisting of 50 million words taken from the *New York Times* over a six month period. She ventures some methodological concerns as the method of measuring adjective-noun co-occurrence is less reliable with less frequent adjectives, for example *damp*. Thus the corpus data was supplemented with examples from the above mentioned dictionaries as well as texts from English Literature available through Project Gutenberg.[20]

Another methodological restriction was that, despite POS tagging, the phrase structure of sentences was not marked and it was therefore not possible to select predicative uses like *Her house was large* (Muehleisen 1997: 65) thus missing out on a large number of co-occurrences. This presented very few problems with the first set of adjectives, *large:small* and *big:little*, but was more relevant with *happy:sad/unhappy*.

The three extremely thorough case studies, which I will not discuss in detail here, reveal interesting patterns in semantic range. They seem to bear out the intuition that better antonyms should have a larger overlapping semantic range than pairs which are only considered peripheral examples of antonymy. Most of the information on semantic range was collected by simply looking at which nouns and kinds of nouns (e.g. physical objects, actors, organisations etc.) these adjectives most commonly modify.

---

[20] www.gutenberg.org (20.03.2010)

Connotational meaning was found to play an important role in semantic range as, for example, the connotational meaning of *little* severely restricts its semantic range. Contrary to its dictionary definitions that propose a wider application, it is, only used in fairly restricted environments, namely when describing 'nouns which name people (e.g. *girl*, *kids*), their personal possessions and pets (*dresses, dog*) or places where they live (*town, village*)' (Muehleisen 1997: 202). Another factor is specificity of meaning: the lexemes *parched* and *arid*, for example, are so specific in meaning that they are largely restricted to describing land which has been affected by the weather.

Frequency, not frequency of co-occurrence, unsurprisingly also plays a role in the application of the concept of semantic range, as the most frequent lexemes have the widest semantic range. This argument could also be constructed the other way around, as the fact that a lexeme can occur with a large number of other lexemes would ensure that it occurs more frequently than one which is more restricted in its application.

A last interesting point to mention in connection with semantic range is the fact that when an opposite is morphologically derived, the semantic range shows a larger overlap with that of the base antonym (*happy:unhappy*) than when the two are morphologically unrelated (*happy:sad*).

Muehleisen concludes her observation by stating that, although frequency of co-occurrence is often cited as the sole decisive factor for 'good' antonymy, there must be a reason for the lexemes to co-occur in the first place and this reason is a large proportion of shared semantic range.

I will now attempt, very broadly and without the finely tuned methodology used by Muehleisen, to exemplify the notion of semantic range using the temperature scale proposed above. Using only websites collected through Google[21] as a corpus, as I have done here, the most frequent of the lexemes above are *hot* (601 million), *cool* (414 million), *cold* (252 million) and *warm* (184 million) and the least frequent are *tepid* (2 million) and *lukewarm* (4 million). *Boiling* (25 million) and *freezing* (31 million) inhabit a middle ground. Assuming that the large number of hits for *cool* is, at least partly, influenced by the use of *cool* as a colloquial term of approval, even these frequency counts support the ranking which was presented earlier and was based wholly on intuition: *hot* and *cold* are the most frequently used lexemes in this list and are, presumably, also the most collocationally versatile. They are followed by *cool* and *warm*, *freezing* and *boiling* and, finally, *tepid* and *lukewarm*

[21] Here, I have used the whole of the web as a corpus without restrictions (12.04.2007). A full analysis with data from the BNC is carried out in 3.3.2.

in descending order. As frequency decreases we also assume a smaller collocational range, although that is only an assumption and not necessarily transferable to other lexemes.[22]

### 2.5.3 The influence of co-occurrence and entrenchment

All the characteristics discussed above are characteristics of a conceptual nature. There are, however, those who argue that ('canonical') antonymy is not determined by a conceptual link but by an associative one (among others Miller et al. 1989, Charles & Miller 1990, Justeson & Katz 1991/1992, Fellbaum 1995). This distinction also has important consequences for assumptions about the structure of the adjectival mental lexicon. Evidence for the influence of frequency of co-occurrence will be discussed in 2.5.3.1 before a somewhat more refined proposal, that of breadth of co-occurrence (Paradis et al. 2007) is illustrated in 2.5.3.2.

#### *2.5.3.1 Frequency of co-occurrence*

It seems, in the first instance, counter-intuitive that what has always been termed a 'paradigmatic lexical relation' should be dependent on a syntagmatic criterion such as frequency of co-occurrence. Nevertheless, this criterion seems to play an important role in determining how strong the relationship between the two members of an antonym pair is. Whether frequency of co-occurrence only has an indirect influence on antonymic strength by increasing the associative strength of the pair which, in addition to all other criteria, makes the pair a 'better' antonym pair overall, or whether it can be considered a factor in its own right remains to be discussed. If the connection between antonyms is indeed an associative one, entrenched by frequent co-occurrence, then the adjectival lexicon is structured completely differently from the lexicons of nouns and verbs, which are very much assumed to be structured on the basis of similarities rather than differences.

With the rise of corpus linguistics, it was discovered that 'canonical' antonyms co-occur at a much higher rate than statistically expected (Justeson & Katz 1991/1992, Mettinger 1994, Fellbaum 1995 among others). This was put forward as the single deciding factor in determining whether an antonym pair was to be considered a 'good' or 'canonical' pair or not. Frequent co-occurrence would lead to a deeper entrenchment of the lexemes in combination, just as it does in

[22] A full analysis of semantic range is not attempted here as the example is only inserted to keep the same scale as for the other criteria.

collocations or idiomatic expressions, regardless of semantic criteria, and therefore to the greater associative strength generally displayed by canonical antonyms.

I will now take a closer look at the two positions regarding frequency of co-occurrence: firstly the position which states that antonymy is a conceptual relation and frequency of co-occurrence is merely another factor which helps to determine whether an antonym pair falls within the category of canonical antonyms; and secondly the opposing view which sees antonymy as a lexical relation between words – a purely associative link.

Starting with the latter perspective, Miller (1990: 242) states that 'antonymy is a lexical relation between word forms, not a semantic relation between word meanings' (cf. also Gross et al. 1989, Charles & Miller 1990 and Gross & Miller 1990). They base this hypothesis on the fact that near-synonyms of a certain lexeme, while semantically opposed, are not considered an antonym pair, or at least not a 'good' match for the original lexeme. *Ascend*, for example, is considered a synonym of *rise* but *ascend:fall* is not considered a good antonym pair and neither is *rise:descend*. The canonical pairs here are *rise:fall* and *ascend:descend*. Therefore, the relation established here is not one between the semantic meaning of the words, as these are identical (or at least very similar), but between the word forms themselves. It is often argued that this associative link between certain word forms can be traced back to a higher than expected rate of co-occurrence of these lexemes. The pairings between the examples above, however, can also be explained by register or the concept of semantic range, introduced in the previous section, as *rise*:*ascend* will presumably differ in semantic range, as will *fall*:*descend*.

This hypothesis, put forward by Charles & Miller (1989), is known as the **co-occurrence hypothesis**. It replaces the **substitutability hypothesis** which claimed that 'the context in which one word occurs in a sentence may call to mind other words syntactically substitutable in that context, so that substitutable words are activated as mentally in close temporal proximity' (Justeson & Katz 1991: 2) and was introduced by Ervin-Tripp (1961, 1963).[23] Charles & Miller 1989 state that the context of any given occurrence of an antonym usually only leaves one member as a possibility and they therefore reject the substitutability hypothesis.

The co-occurrence hypothesis was then supported and extended by findings presented in Justeson & Katz 1991. They state that their research 'confirms that, for all adjectives frequent enough to judge, antonymous adjectives do co-occur within

---

[23] The substitutability hypothesis is only mentioned for completeness' sake and will not be discussed further as it has been proven to be of little impact in recent discussion. For a more detailed account see Charles & Miller 1989 as well as Ervin-Tripp 1961 and 1963.

the same sentence much more often than is expected by chance' (Justeson & Katz 1991: 3). They further state that it is not co-occurrence alone that accounts for the 'training effect', but co-occurrence in similar sentential structures where often the only change is the substitution of one antonym of a pair for the other (cf. Justeson & Katz 1991: 3 and also Jones 2002).

Finally, Justeson & Katz (1991: 13) reiterate the theory they propose: 'co-occurrence takes place via substitution, substitution yields antonymy alignment, and alignment leads to association' and thus combine the co-occurrence hypothesis with a more precise version of the substitutability hypothesis to account for the strong associative bonds observed in antonymy. They conclude by calling antonymy a lexical-semantic relation but stressing that the lexical relation is the one which makes antonym pairs canonical.

Fellbaum 1995 extends the co-occurrence hypothesis, which had until then only been applied to adjectival antonym pairs, to nouns and verbs, and states the following hypothesis in analogy to the co-occurrence hypothesis proposed by Charles and Miller (1989):

> **H1: The Co-Occurrence Hypothesis for Semantically Opposed Nouns and Verbs**
> Semantically opposed nouns tend to be used together in word pairs in the same sentences, as do semantically opposed verbs. (Fellbaum 1995: 286)

Fellbaum tests this hypothesis on three concept pairs (*lose:gain*, *start:finish* and *increase:decrease*) in the Brown Corpus (1 million words; Francis & Kucera 1982). These three pairs were chosen for representability of semantic opposition and frequent occurrence. The data extracted from the corpus supports the hypothesis proposed above and only differs from the adjectival findings by Justeson & Katz in one significant way: verbal and nominal pairs tend not to occur in parallel syntactic constructions, but verbs frequently have different subjects and nouns can differ in number or can be part of different compounds (cf. Fellbaum 1995: 287).

The second hypothesis put forward by Fellbaum extends the co-occurrence phenomenon across word classes:

> **H2: The Co-Occurrence Hypothesis for Semantically Opposed Concepts**
> Regardless of their syntactic category, words expressing semantically opposed concepts tend to be used together in the same sentences. (Fellbaum 1995: 289)

This hypothesis is also borne out in the data collected from the Brown Corpus. However, the pairs tested appear in different syntactic environments and therefore, once again, the substitutability hypothesis cannot be upheld. Fellbaum further proposes a number of syntactic frames in which opposed concepts frequently occur, a notion which is taken up in Jones et al. 2007 and is discussed in detail in 2.5.3.2.

The other perspective on antonymy, that of a conceptual relation, does acknowledge frequency of co-occurrence as a factor contributing to antonymy but does not see this as the all-determining criterion. One of the main queries is, as Murphy and Andrew (1993: 304) put it, 'if Miller et al. (1990) are correct in asserting that antonymy is a kind of lexical association, then the question arises as to how opposites become associated'. Murphy and Andrew are furthermore sceptical about the power of co-occurrence as this explanation does not seem to do justice to the close semantic relation displayed by all antonyms, but especially canonical opposition. This view is also shared by, among others, Murphy 2003, Cruse 1990, Cruse & Togia 1995 as well as by the present author.

The issue arises that 'if the antonym relation can easily be explained by conceptual relations, there seems to be little reason to require a special kind of lexical association in addition' (Murphy & Andrew 1993: 304). The present account also takes stylistic factors, as well as the question of register and semantic range, into account, which can explain some of the data which has been attributed to co-occurrence only.

Murphy and Andrew put forward the hypothesis that 'antonyms are defined by their conceptual opposition on a single dimension; any other kind of difference between the two words weakens the sense of antonymy.' (1993: 306). These differences might be of a morphological or stylistic nature or the antonymous relation might be obscured by added semantic content (as discussed in 3.3.1.2). They conclude that lexical relations may be influenced by co-occurrence but that it is generally possible to determine the antonym of a certain lexeme by the principle of minimal difference applied to semantic as well as stylistic and morphological properties of the lexeme in question. The best antonyms will generally only differ in **one** respect, which will usually be a salient conceptual one. This also accounts for the fact that some lexical relations can be reliably constructed without having occurred before and must therefore be to some extent derivable since otherwise an on-line construction would not be possible.

Returning, one last time, to the example of the TEMPERATURE scale, it is clear that good antonyms co-occur more frequently than less good ones, which can be illustrated by further Google data: *hot:cold* (89 million), *tepid:lukewarm* (97,600), *freezing:boiling* (1 million) and *warm:cool* (40 million). Again bearing in mind the idiosyncrasies in Google data, these figures nevertheless show support for the intuitive judgement made above when choosing *hot:cold* as the best antonym pair.

*2.5.3.2 Breadth of co-occurrence*

In a recent study, Paradis et al. (2007) investigated the influence of frequency of co-occurrence on antonym judgements more closely on the basis of a hypothesis which claimed that it is not simple co-occurrence which makes antonyms special but co-occurrence within certain syntactic frames (cf. Figure 2.8). A large-scale corpus study of ten adjectives and their opposites (as determined in elicitation experiments (Paradis et al. 2009) was conducted based on the premise that antonyms 'have a tendency to favour certain lexico-grammatical constructions in discourse' (Paradis et al. 2007: 1). The two goals which pertain to the matter in hand were the question whether it is feasible to 'use a series of lexico-grammatical constructions as diagnostic of antonym' (2007: 1) and whether the fidelity with which certain pairs co-occur in these constructions can be used as a measurement of their antonymic strength.

Murphy 2006 states that the frames in Figure 2.8 are inherently contrastive and thus the most semantically incompatible lexemes will be the ones most frequently seen in these frames. Thus, the investigation of these syntactic frames can be used as an 'antonym-discovery methodology' (Paradis et al. 2007) since, if the frame is entered into a corpus with the X-position filled with an adjective (e.g. *hot*), the results for the Y-position should be antonymically related to *hot* (e.g. *cold, cool, freezing*) and the more frequently an item from the Y-list comes up, the stronger the antonymic relation is perceived to be. Thus, if the above example were to be entered in a corpus search machine, *cold* should head the list of results generated for Y.

| | |
|---|---|
| X and Y alike | from X to Y |
| between X and Y | X versus Y |
| both X and Y | whether X or Y |
| either X or Y | |

**Figure 2.8** *Lexico-grammatical constructions used by Paradis et al. 2007*

The results of the corpus investigation carried out by Paradis et al. (2007) were largely consistent with Paradis et al.'s later elicitation task results (2009). However, some of the Y-results in some frames were not antonyms of the lexeme in the X position (2007: 9); for example, *boring* triggered *dull* as well as antonymically related adjectives like *interesting* or *exciting*. Co-ordinating constructions like *either X or Y* and *both X and Y* thus seem to be able to be used with more than one meaning and/or discourse function.

Paradis et al. (2007) claim that breadth of co-occurrence, i.e. the co-occurrence of two lexemes in a number of inherently contrastive syntactic frames, is a more reliable indicator of antonymy than simple frequency of co-occurrence. This does seem to be the case, as simple co-occurrence data merely reflects associative strength and therefore also includes collocations, unique binominals and other frequently co-occurring items which are not opposites at all (for a more detailed discussion see Chapter 6). Thus, breadth of co-occurrence does indeed appear to be a diagnostic which is more appropriate as well as more reliable in the detection of antonymy (2007: 9).

It is, however, far from fool-proof and this approach has not been able to determine whether the occurrence of antonymic word pairs in the investigated syntactic frames is due to the fact that the construction invites a contrasting pair (in which case the word pair would have to be antonymic as a prerequisite for use in these constructions) or whether the constructions take any type of word pair and imbue it with a certain degree of opposition. This goes back to the question whether antonyms co-occur frequently because they are antonymic or whether their frequent appearance together has entrenched the connection in our minds and thus contributes to the pair's oppositeness (cf. Murphy & Andrews 1993, Murphy 2003, 2006, Paradis et al. 2006).

There is no doubt that a corpus search of the above-mentioned frames is a tremendously powerful tool for the detection of antonyms, much more so than a simple co-occurrence search, and that the data collected through the former approach is much more homogeneous. However, the question whether the distribution of antonyms in the lexico-grammatical structures is a consequence of opposition or a criterion for it is still very much under debate (see Chapter 6 for further discussion).

Both frequency of co-occurrence and breadth of co-occurrence seem to play a role in determining the antonymic strength of a pair of lexemes but how precisely these factors influence speaker perception and judgements of lexical opposition is yet to be determined.

## 2.6 Antonymy as a gradient phenomenon

At the end of this discussion of factors influencing the canonicity of an antonym, having followed the temperature scale pairs through the whole discussion, we can now reliably conclude that *hot:cold* is the most canonical of the four opposite pairs investigated. It fulfils all the criteria proposed by Cruse: the two lexemes seem to share a large proportion of their individual semantic range, they are similar in style and register, and co-occur frequently in sentences, especially in certain lexico-grammatical frames.

In a follow-up paper to Justeson & Katz 1991 (where they claim that the varying strength of antonymic relations in speaker judgements depends largely on the lexical criterion, namely frequency of co-occurrence), the authors investigate the co-occurrence phenomenon more closely and discover that 'almost all semantically related adjectives co-occur in significant numbers' (Justeson & Katz 1992: 184). Therefore the essence of the antonymic relation lies irrevocably in the semantic component, but frequent co-occurrence is likely to play a large part in associative strength decisions. This is one of the main arguments this thesis will investigate, by taking into account those peripheral phenomena of opposition which have, to date, not been considered of great interest.

Another conclusion which can be drawn here is that antonymy is by no means an either-or relation but that, as Justeson & Katz (1992: 182) put it, 'adjectives may be more or less antonymous rather than simply antonymous or not antonymous.' This argument, proposing an 'antonymicity continuum', has found further support in more recent work (Murphy 2003, Paradis et al. 2007, Murphy & Jones 2005, Paradis et al. 2009 and Murphy et al. 2009) and forms one of the underlying hypotheses of the current research.

In addition to providing further evidence for antonymy as a gradient phenomenon, the empirical data presented and analysed in the following chapters will lead to a catalogue of factors determining an antonym pair's place on the proposed antonymicity continuum and thus provide a clearer picture of the factors influencing the lexico-conceptual relation of lexical opposition and the internal structure of antonym categories.

# 3. Associative and antonymic strength

The opposite of a *king*, I'm sure,
Is someone humble and obscure –
A *peasant*, or some *wretched soul*
Who begs through life with staff and bowl
Another opposite's the *queen*,
If she is quarrelsome and mean.
(Wilbur 2004: 507)

This chapter will analyse the results of the goodness-of-exemplar (GOE)-rating questionnaire which forms the basis of the psycholinguistic experiments discussed in Chapter 4. The GOE-rating questionnaire was designed in keeping with Rosch's original goodness-of-exemplar questionnaire (cf. Rosch 1975). The principle remains the same despite the fact that the objective and the type of judgement required from the participants in this study are slightly different; for ease of reading it will nevertheless be referred to as a goodness-of-exemplar rating task. Whereas regular goodness-of-exemplar tasks require the participant to draw on their conceptual knowledge of a certain category (e.g. FURNITURE or BIRDS), in this case the informants are asked to make a meta-judgement about the relationships between certain lexemes and thus categories. Therefore, the judgements are considered less immediate and less automatic and they furthermore depend, to a certain extent, on the individual's concept of opposition. However, despite these issues, which could be seen as drawbacks, this technique can nevertheless be used as an effective and reliable method of assessing antonymic strength and goodness of antonymy between members of a variety of word pairs. It allows a focus on specific pairs as well as sequencing; both of these aspects could be lost when a free association task is used, and an antonym elicitation task would also involve judgements at a meta-level and thus have some of the same criticisms levelled against it as the GOE-rating task. Several studies have successfully used judgement tasks as a method of determining degrees of similarity and difference between certain lexical items (cf. Murphy & Andrews 1993, Charles et al. 1994) to calibrate stimuli for behavioural experiments.

Two different measures will initially be discussed separately and then in relation to each other: the GOE-rating results, and measures taken from the Edinburgh Word Association Thesaurus (EAT).[1] The former is considered a measure of antonymic strength whereas the latter, like frequency of co-occurrence, is considered solely a predictor of associative strength (which includes relations

[1] http://www.eat.rl.ac.uk (last accessed May 2010)

such as synonymy and collocation). Data collected from both of the above sources will be compared to frequency of co-occurrence data taken from the British National Corpus (BNC). The focus of the analysis in this chapter, however, is on the data generated by the GOE-rating questionnaire, which was designed primarily to highlight potential areas for closer investigation and therefore deals with a fairly large number of antonym pairs. One of the main questions is whether the measures of associative strength (EAT and BNC data) serve as a useful and reliable predictor of antonymic strength or whether other factors play a more central role in antonym judgements. Some of the factors considered stem from previous research (cf. 2.5) and will be re-examined with the help of the GOE-rating data, while others are introduced to explain phenomena observed in the data, which cannot yet be accounted for by the existing catalogue of criteria for good antonymy.

Not all pairs included in the questionnaire will be considered individually but, once the design, selection criteria and methodology of the questionnaire have been discussed, a number of particularly interesting cases among the results will be examined in greater detail. These cases will also be revisited in the following chapters (Chapter 4 and 6) where the additional results from lexical decision tasks will be added to the data considered in the present chapter and in Chapter 5 where the English data is compared to the German data. A summary of features which influence antonymic strength and their applicability will conclude this chapter (3.5).

## 3.1 Selection criteria for antonym pairs

This section aims to illustrate the selection criteria used in determining the 210 word pairs used in the questionnaire. Some of the criteria are lexical (e.g. word class, morphological relatedness) and others are based purely on previous research in antonymy. Rather than including only canonical pairs, or canonical pairs with indirect/peripheral counterparts (e.g. *wet:dry* and *soggy:parched*), a slightly different approach was used which also includes less conventional pairings (*tea:coffee, landline:mobile* or *credit:debit*) which conform to certain criteria (cf. 2.5.1 & 2.5.2) as well as relations such as converses (*buy:sell*). Furthermore, some pairs form clusters around a certain canonical pair (cf. 3.4.1). A visual summary of all pairs included in the questionnaire can be found in Figure 3.2 and a full list in Appendix 2. The examples discussed in greater detail in this chapter are those of the adjectival lexical fields of *hot:cold*, *good:bad* and *big:little/large:small,* as well as nominal and verbal converses and the case of the complementary pair *male:female*.

### 3.1.1 Word class

Naturally, the best examples of antonymy are adjectives and therefore the questionnaire shows a clear bias towards adjectival pairs (57.21%). However, due to the inclusion of converses (see below), there is also a comparatively large number of nouns (22.12%) and verbs (11.54%), including a number of pairs which could be either nominal or verbal. Prepositions and adverbs were also included but in much smaller numbers than any of the groups above (4.33% and 2.4% respectively).

There are no pairs consisting of members of different word classes, despite the fact that these pairs also co-occur more frequently than chance would predict (e.g. *life:dead*; cf. Fellbaum 1995), as this does not constitute the focus of the present study. Pairs in which one member could be classified as belonging to two different word classes were classed according to the other member of the pair as it is assumed that informants are more likely to interpret the items as members of the same word class.

### 3.1.2 Type of antonymic relation

An important factor in choosing antonym pairs for this study was type of antonym. According to Lyons 1977 and Cruse 1986, as discussed in 2.1.1 and 2.1.2 respectively, there are a number of different types of antonym pairs. It is not clear whether these classifications, which rely very strongly on the logical and syntactic properties of any given opposite pair, do in fact influence judgements of antonymy. The pairs which are usually deemed to be the 'best' examples by researchers are the ones Cruse and Lyons call antonyms in the strict sense: binary, gradable, adjectival opposite pairs (*cold:hot*). However, because it is unclear whether these pairs are indeed more antonymous than, for example, non-scalar pairs, all types of pairs were included in the questionnaire. The sections below explain the selection criteria in this category in greater detail.

*3.1.2.1 Gradable opposite pairs (antonyms)*

The questionnaire contains 69 gradable antonym pairs, some of which are simply individual pairs which were considered interesting either in meaning or in other properties (e.g. *dirty:clean* (privative antonym); *fast:slow* (highly canonical); *drunk:sober* (highly gradable)). The remainder of the gradable antonyms are clusters based on highly canonical antonym pairs (*hot:cold*, *good:bad* and *big:little/large:small*). These clusters consist of what are often called 'direct' and 'indirect' antonym pairs (cf. for example Gross et al. 1989), where the members of indirect pairs are selected from among the synonyms of the base pair. Figure 3.1 will help to illustrate and explain the selection process used in the present research for the lexemes related to the base pair *hot:cold*.

The bold blue arrow indicates the temperature scale from *cold* on the left to *hot* on the right. The green arrows represent connections between more central (and more conventional) antonym pairs in this field whereas the orange arrows stand for less conventional connections between synonyms of *hot* and *cold*. The pairs linked by red arrows are pairs in which the lexemes are not linked to members of the same lexical field but to members of a field which is related through one of the other meanings of *hot* ('spicy') or *cold* ('unfriendly').

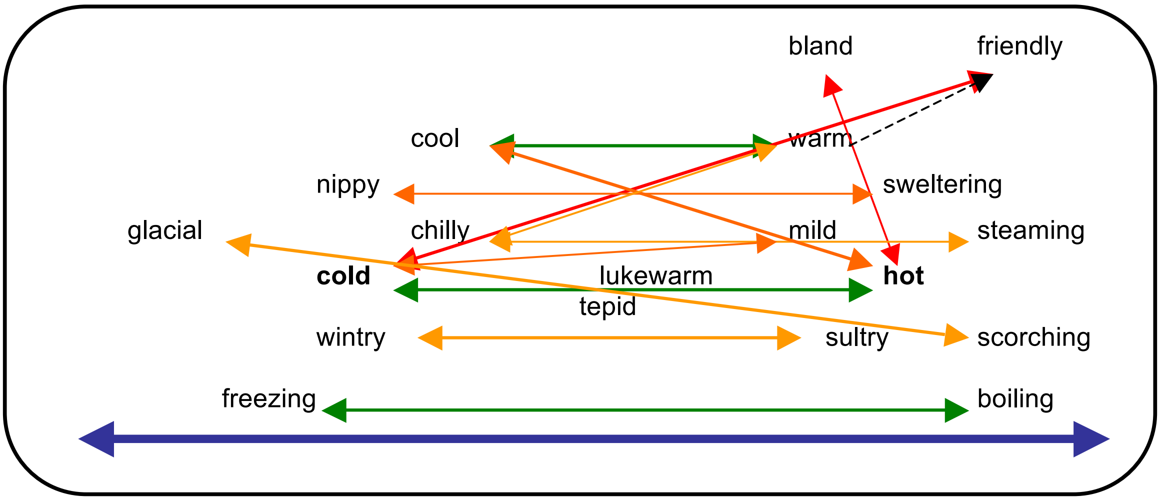



**Figure 3.1** *Cluster around base pair* hot:cold *(example of the selection process of gradable pairs for this study)*

The three pairs linked by green arrows (*hot:cold, warm:cool* and *boiling:freezing*) are all symmetrically distributed along the temperature scale (equidistant from the mid-point of the scale) and, together, cover a significant proportion of the scale. Out of these, *hot:cold* would be expected to score highest, followed by *warm:cool* and then *boiling:freezing*, since *hot:cold* marks the 'standard' and most salient distance from the midpoint of the scale whereas the other two pairs denote more narrowly defined concepts of temperature at more extreme points of the scale.

The orange links connect words to form pairs which either violate one of the criteria for good antonyms (cf. 2.5) or are simply composed of randomly selected synonyms of the base pair. The pairs which contain a deliberate criterion violation, for example a mis-match in distance from the midpoint of the temperature scale (e.g. *cool:hot, warm:cold*), are expected to score slightly lower than those which conform to the criteria but higher than pairs such as *glacial:scorching* which presumably do not have enough shared semantic material (and do not co-occur frequently enough) to provide the same level of entrenchment that is found in some of the other pairs. In contrast to other studies (e.g. Charles et al. 1994) which only consider permutations which include either member of the base pair, pairs which do not contain either of the members of the base pair but operate on the same scale are also included. This approach was used to select a number of pairs for each of the proposed clusters of gradable opposites.

#### *3.1.2.2 Non-gradable opposite pairs (complementaries)*

The non-gradable pairs included in this study could initially be expected to score extremely high on the seven-point scale used in this questionnaire. Prominent antonyms in this category are *true:false*, *male:female* and *right:wrong*. The method used with gradable pairs was repeated with these complementaries and an array of related pairs was chosen.

In the case of *male:female*, the related lexemes focussed mainly on violations of the purity of opposition criterion as the scalar properties do not apply here and the chosen pairs are largely nominal. These pairs include the following: *mother:father, cow:bull, king:queen*, and *masculine:feminine*.

The clusters around *true:false* and *right:wrong* included a range of other adjectival synonyms of either member of the pair. Some of these were paired with a prefixed partner which did not have the same stem: *exact:inaccurate, right:incorrect*. Other pairs included here were *precise:imprecise*, which seems to be a borderline case where gradability is concerned (cf. privative opposites in 2.1.2), and *suitable:wrong*, as well as the ‘correct’ pairings for both morphologically mismatched pairs *unsuitable:suitable* and *accurate:inaccurate*. This cluster will not be discussed in detail in this chapter but will be fully analysed in the discussion of the importance of morphological relatedness in Chapter 6.

#### *3.1.2.3 Converses*

Another group, which is included despite often not being classed as lexical opposition, is that of converses, which is discussed in greater detail in section 3.4.2. The inclusion of converses (e.g. *buy:sell, come:go*) provided a greater number of verbal and nominal opposites which are not morphologically related. This is an attempt to extend the research carried out so far to less central members of the category of lexical opposition and to discern whether the theories posited for the smaller number of canonical opposites still hold when a larger number of peripheral antonyms is taken into account.

Converses represent, in themselves, a very interesting category which is often said to not be part of antonymy proper as it exhibits several distinct characteristics (cf. 2.1.2 & 3.4.2). However, many converses are readily referred to as opposites in everyday language, and therefore can also be expected to be judged as opposites in the GOE-rating task.

The questionnaire contains ten pairs of verbal converses and nine pairs of nominal converses. All of these pairs will be discussed in detail in section 3.4.2 as a comparison between the frequency data on the one hand and the GOE-rating data on the other reveals some discrepancies which merit fuller discussion.

#### *3.1.2.4 Spatial (directional) opposites*

Directional opposition, as defined by Cruse (1986), would contain reversives (*pack:unpack*) and restituitives (*damage:repair*) as well as converses and the sub-category which I will refer to as spatial antonyms (*in:out, up:down*): those lexemes which help us anchor reference points in spatial perception. Taking cognitive theories into account (Lakoff & Johnson 1980, Lakoff 1987), these pairs should be particularly salient and would therefore be expected to score reasonably high in the GOE-rating as well as the associative strength measures from the EAT.

The following ten pairs were included in the study: *in:out, up:down, last:first, together:apart, next to:opposite, top:bottom, over:under, far:near, behind:in front* and *right:left*. Most of these pairs, with the possible exception of *next to:opposite, behind:in front* and *together:apart,* are very strongly conventionalised as pairs and also serve as scaffolding for many metaphoric expressions which are built around the spatial structures these pairs represent (cf. for example Lakoff & Johnson 1980). The group of directional opposites is discussed contrastively (English-German) in 5.1.2.

### 3.1.3 Morphological relatedness

The number of morphologically related opposite pairs in this study is much smaller than the number of unrelated pairs as relatedness plays an important role in antonym judgement (cf. for example Murphy & Andrews 1993) and antonyms which share a common root are generally considered better pairs than ones which do not (e.g. *ascend:descend* vs. *rise:descend*). Relatedness sometimes even influences participants to make the 'wrong' decision about a word pair (e.g. *flammable:inflammable, helpful:helpless*), thus attributing antonym status on the basis of shared morphological material despite the fact that the two lexemes do not stand in any antonymic relation to each other.

There are 23 morphologically related pairs in this study and an attempt was made to include all common negative prefixes (*un-, in-*, *dis-*). Initially *non-* was also included, but as it is used frequently in ad-hoc constructions, it was later discarded because of doubts as to whether it would contribute to the results of the study. Generally, the morphologically related pairs are expected to receive high or very high scores in the GOE-rating; this means they would be good to excellent examples of antonymy (cf. 3.1.4 for groupings of antonymy).

Some pairs are made up of one lexeme with a negative prefix and an opposing lexeme which is not morphologically related (e.g. *exact:imprecise, good:disobedient*). Despite the fact that there is only a relatively small number of these lexemes, they should be indicative of the importance of morphological relatedness (especially of prefixation) in antonym judgements.

While the results of the GOE-ratings of pairs which share morphological material are relatively close together, much more variation is expected from the unrelated pairs as their relationship is not as transparent and immediately obvious, but relies solely on other, less transparent criteria to determine the strength of their relation. One could almost call the morphologically related pairs a positive control as they give an indication of a 'typical' antonym score on a GOE-rating scale. Morphological relatedness will be discussed in its own right in 6.1 as well as in Chapter 5 (cf. 5.1.1) as part of the contrastive analysis of the English and German data.

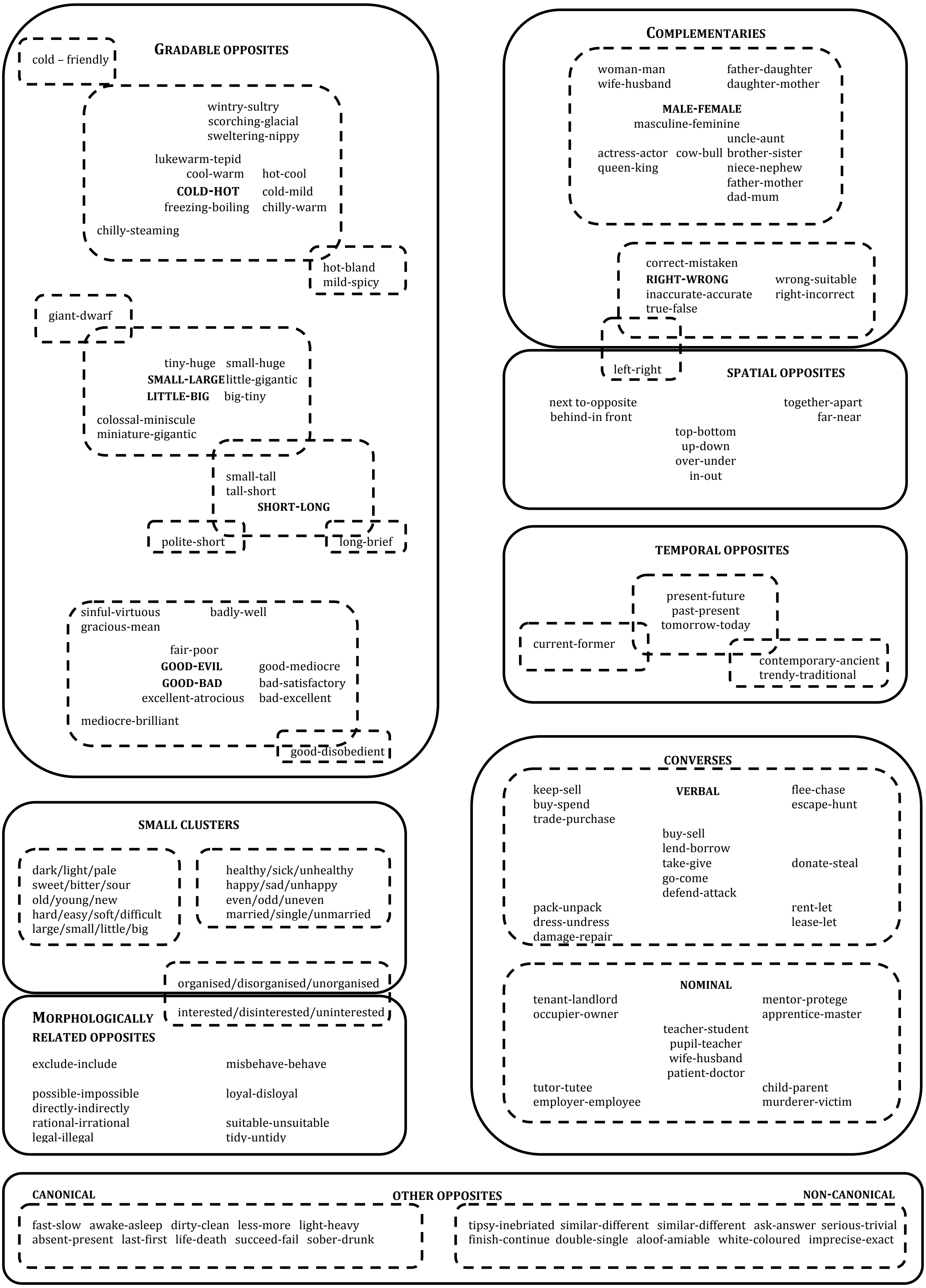



**Figure 3.2** *Opposite pairs included in the GOE-rating questionnaire*

### 3.1.4 Control Items

Various pairs were inserted as control items, all of which display some measure of relatedness and associative strength, e.g. meronyms (*book-page, hand-finger*), synonyms (*yell-call, sad-unhappy*), co-hyponyms (*mug-cup, dog-cat*), hyponyms (*animal-dog, flower-rose*), as well as lexical items which simply have a high associative score or frequency of co-occurrence (*fish-fowl, chicken-egg, hot-humid*) or which are structurally similar to antonym pairs (*helpful-helpless, flammable-inflammable*) due to their morphological properties.

To ensure a balance of scores across the whole scale of one to seven, all pairs used in the questionnaire were divided into four groups depending on goodness-of-example of antonymy as judged by the researcher. Each of the four groups comprised roughly 50 pairs (groups I and II were slightly larger) and some items in group IV were used twice (cf. 3.2). This approach resulted in a distribution of 25% for each GOE-rating band: group I contained antonyms judged to be excellent examples of the category; group II pairs which were slightly less convincing examples and would usually violate one of the criteria illustrated in 2.5; group III contained items which could still be seen as antonymous but which violated several crucial criteria. Lastly, group IV contained the control items mentioned above which are not judged antonymous at all but nevertheless scored reasonably high on associative strength measures. This was intended to result in a relatively even distribution of scores and to minimise a ceiling effect or training of the participants towards one particular end of the scale. Initially, on the scale of one to seven, the four bands are set as follows: Group I 1-1.79 (64 items), group II 1.8-2.99 (41 items), group III 3-4.99 (57 items) and group IV 5-7 (48 items). As the number of pairs in each of the groups illustrates, the results show a similar pattern to that expected. Group I is slightly larger than was originally intended but, that aside, the distribution is relatively even.

## 3.2 Methodology

A pilot questionnaire was conducted and several changes were made to improve the effectiveness of the questionnaire. One of the major changes was the decision to split it in half (104 and 106 pairs), since rating 210 pairs in one questionnaire was a fairly arduous task for the participants. The second major change was to increase the number of pairs which would score at the bottom end of the scale (7) as the pilot questionnaire revealed a considerable bias towards the higher end of the scale (1).

Furthermore, some of the pairs included in the pilot study are no longer present in the questionnaires discussed below as the selection criteria were changed slightly in the time between the two questionnaires. One of the consequences of the decision to work with certain types of antonym pairs was the omission of some of the canonical pairs, which are usually included in antonym studies (e.g. *fast:slow*, *black:white*). The questionnaire nevertheless still contains a significant number of highly canonical pairs.

The changes outlined above resulted in the construction of four questionnaires as the order of the two members also had to be reversed for each pair. Questionnaire I-1 contains 104 pairs, half of which appear in their prototypical order (insofar as a prototypical order can be determined; e.g. *up:down*) and the other in the reverse. Questionnaire I-2 contains the same pairs but in reverse order. Exactly the same applies to Questionnaires II-1 and II-2 which contain 106 word pairs each (cf. Appendices 1 and 2).

The questionnaires consisted of a short introduction and instruction section, a personal data section - name, age, gender, first language, hometown and variety of English – which had to be filled in before the questionnaire could be completed, and the main section of 104 (or 106) word pairs to be rated on a 7-point scale. At the end of the questionnaire, participants were given the opportunity to comment on the task as a whole and mention any specific issues concerning a certain pair, but not many participants took this opportunity and the comments mainly related to the fact that the decisions were sometimes not as easy as anticipated.

### 3.2.1 Task and procedure

The four questionnaires were created with a web-based questionnaire tool,[2] which allows the user to create questionnaires of any size; it furthermore collates the responses and creates a summary, as well as separate records of individual responses which can easily be browsed. The order of pairs in each of the questionnaires was also randomised automatically for each participant.

Web links to the four questionnaires were sent out by e-mail, mainly to students and staff at the University of Oxford, with a letter explaining what the questionnaire would be used for and how to complete it as well as the information

[2] www.surveymonkey.com (accessed May 2008 - April 2010)

that ethics approval had been obtained. In addition to this, the recipients of the initial e-mail were asked to pass the letter and web link on to other native speakers of English.

The task itself was a straightforward goodness-of-exemplar rating task (cf. Rosch 1973, 1975, 1978) during which the participants were asked to rank each pair on a Lickert scale from 1-7 with 1 being excellent and 7 very poor. The instructions furthermore asked participants to fill the questionnaire in swiftly and not to go back to change their responses once they had made a decision, to prevent over-rationalisation on the part of the participants. However, as the questionnaires were not completed in a controlled environment with the researcher present, it is impossible to ascertain whether the informants kept to this instruction. Some participants did report that they found this difficult as on occasion they wanted to change their mind retrospectively.

### 3.2.2 Participants

160 respondents participated in the questionnaire study; 40 completed each of the separate sub-questionnaires. During data collection, some questionnaires which were filled in incorrectly (e.g. the respondent took 7 as excellent instead of 1) were immediately discarded. As the collection method was an online questionnaire, it was possible to simply delete the unusable questionnaires and continue with the data collection until the required number was reached.

The only formal criterion for the task was that the informants had to be native speakers of English. Variety of English was deemed not to have a great influence on the GOE-judgement of antonym pairs and respondents were mainly speakers of British, Canadian and American English with a few who were speakers of New Zealand and Australian English.

The age of the participants ranged from 18 to 60, the vast majority being between 20 and 35. Age could prove to be an interesting factor, especially concerning some of the less conventional antonym pairs included in the study (*interested:uninterested* vs. *interested:disinterested; landline:mobile*), but it was initially not considered a variable to be included in the present study.

Lastly, all speakers were educated to A-level standard, and almost all had a university degree, with many in postgraduate education. However, education was not considered a factor to be taken into account at this stage since, even though this information was collected with other personal data, it does not seem to have any bearing on the results.

## 3.3 Measures of associative strength

This section introduces the two measures used to supplement the data gained from the GOE-rating questionnaire. The data from the Edinburgh Associative Thesaurus gives information about the general associative strength of the lexical items, regardless of antonym status, whereas the frequency data from the British National Corpus provides a baseline against which to compare the GOE-rating results.

### 3.3.1 The Edinburgh Word Association Thesaurus (EAT)

The EAT provides information on associative strength through data gained from an elicitation task. It can be searched by stimulus or response word. In the former case the output will be a list of words which were elicited for the given stimulus word ordered by number of responses (100 participants contributed to this resource). If it is searched by response word, the results will list all the lexemes the word in question was used as a response for and the number of times it was mentioned.

These measures give a useful indication of associative strength outside the context of antonymy. There are many words (mainly adjectives) among those included in the questionnaire for which the antonym is indeed the primary response (e.g. *good* elicited *bad* 78 times, but *evil* only twice). These results are similar to those described by Deese (1964/65) in early word association experiments. In some cases, however, there is a stronger synonym or otherwise related lexeme (meronym, hyperonym, collocate) which takes the place of the most strongly associated lexeme.

The data collected from the EAT will be compared to the questionnaire results in this chapter but they will also be used for comparison with the data from the lexical decision tasks in Chapter 4. Cases such as *good:bad* and *good:evil*, in which a lexeme has two possible antonym ‘partners’, will also be looked at separately (Chapter 4.2). Here, the EAT data seems to foreshadow a stronger connection between *good:bad* than between *good:evil*. This can presumably be attributed to a difference in semantic range and therefore in frequency of co-occurrence. However, cases where one lexeme has two conventional antonymic partners will be discussed in greater detail in Chapter 4.2.

### 3.3.2 British National Corpus frequency data

As antonymy is a paradigmatic lexical relation, it might seem somewhat counterintuitive to use frequency of co-occurrence as one of the measures of antonymic strength. However, as discussed in 2.2 and 2.5.3.1, it has often been observed that antonyms occur in very close proximity to each other in both spoken and written language and that this is generally considered one of the most reliable indicators of antonymic strength (cf. Justeson & Katz 1991, Fellbaum 1995, Jones 2002, Paradis et al. 2007).

Individual frequencies for all lexemes in the questionnaire were extracted from the British National Corpus (BNC World) using the BNC concordancer Sara98. The British National Corpus contains over 100 million words of written (90%) and spoken (10%) English from a wide variety of sources. It was compiled in the 1990s and is thus still reasonably current.[3]

The co-occurrence data was collected for both lexeme1-lexeme2 and lexeme2-lexeme1 combinations to allow for a later analysis of sequencing preferences among the pairs. The total number of co-occurrences within a span of 5 words was used to calculate t-scores and mutual information as well as the expected/observed ratios, the latter two measures being better indicators of the relevance of the co-occurrence data for small measures. The complete list of co-occurrence data can be found in Appendix 3.

The statistical method used on the data from the British National Corpus was the calculation of t-scores which is the measure most often referred to in the discussion in the present study. Despite the fact that t-scores, which indicate whether the co-occurrence observed in the data is likely to be significant, are less reliable with small numbers and percentages, this method seems to work satisfactorily, especially in conjunction with mutual information and the observed/expected ratio which were also calculated. A t-score of 2.0 is generally seen as the boundary between chance co-occurrence and significant co-occurrence and the higher the t-score, the less likely it is that the co-occurrence of the lexemes is due to chance or other external factors (cf. Appendix 14 (explanation of measures)).

[3] Information from www.natcorp.ox.ac.uk

## 3.4 Antonymic strength: discussion of GOE-rating results

This section will discuss the results of the GOE-rating questionnaires and attempt to determine which factors influence antonym judgements and whether high associative strength is a reliable predictor for high antonymic strength. After a brief discussion of the overall results, three particular cases will be examined in greater detail. Frequency data from the BNC is used as comparison as, if antonymy is indeed determined by frequency of co-occurrence alone, the higher the frequency, the better the antonym pair. Thus, the t-scores should be directly reflected in the questionnaire data.

The examples which have been selected for closer scrutiny are either parts of the questionnaire which were designed to judge the impact of Cruse's (1986) criteria for good antonyms (3.4.1) or examples where a discrepancy can be seen between the frequency results and the questionnaire data (3.4.2 & 3.4.3).

### 3.4.1 Scalar antonyms

Three word fields of gradable, scalar antonyms will be discussed in this section. As *hot:cold* has already been introduced above, the lexemes will simply be listed together with their frequencies and they will be discussed in turn. The second field is that of *good:bad*; this pair displays some slightly different syntactic and logical criteria and is placed by Cruse (1986) in the category of equipollent antonymy. The third, and final, cluster discussed in this section is that around the base pairs *small:large* and *little:big*.

The aim of this section is to investigate whether the pairs were rated as would be expected due to the frequency of co-occurrence data, as well as to examine the consequences of the systematic violations of criteria for good antonymy.

#### *3.4.1.1 The* TEMPERATURE *scale:* HOT:COLD

*Hot:cold* is an example of antonymy proper in the Lyonsian sense, as it is a gradable, adjectival antonym pair. Cruse (1986) calls this subtype of antonymy equipollent antonymy as it contains, conceptually at least, two separate scales which meet at a midpoint (cf. 2.3.1). The pairs along the temperature scale which were chosen for inclusion in the questionnaire are listed in Table 3.1 below and their relationships are illustrated in Figure 3.1 (cf. 3.1.2.1). The table provides the results from the GOE-rating as well as the t-score of the frequency of co-occurrence analysis and the two EAT scores of each pair. The GOE-rating score is the mean of the scores of both pairs (i.e. *hot:cold* and *cold:hot*). Those words which displayed significant differences between the scores of the two pairs will be discussed below. A complete list of results can be found in Appendix 3.

In this section all the above pairs will be discussed in turn and the results from the three sources of data will be compared with the expected results, taking into account the criteria put forward in Chapter 2.5. The more central pairs will be discussed first before more peripheral examples are considered which do not meet some of the following criteria or do so only in a modified form:

- distance from the midpoint of the scale
- non-propositional meaning
- semantic range
- morphological closeness.

Furthermore, the expected ranking, which assumes frequency of co-occurrence to be the single deciding factor, will be contrasted with the actual ranking provided by the GOE-rating questionnaire and the EAT scores. This will then allow a discussion about possible additional factors which need to be taken into account in the judgement of even this most central type of lexical opposition.

The first pairs which will be discussed in greater detail are the three pairs which are the most central on the TEMPERATURE scale and differ only in degree: *hot:cold*, *warm:cool* and *boiling:freezing*. These three pairs all adhere to the criteria above as they are equidistant from the midpoint of the TEMPERATURE scale, have similar non-propositional meaning and semantic range and none of them are morphologically related.

*Hot:cold*, as expected, is one of the highest-scoring pairs in the whole questionnaire and displays all the properties expected from a ‘canonical’ antonym. The t-score is extremely high (20.32) and both EAT scores show a high associative link even under non-antonym conditions. These are the measures which will serve as a reference point for the pairs discussed in the remainder of this section. *Warm:cool* and *boiling:freezing* are intuitively considered less central as their range is more restricted, which could, in theory, lead to a lower t-score as, if a lexeme is bound to certain contexts, this would limit the chances of the two lexemes appearing together. However, as the semantic ranges of these lexemes still largely overlap, especially in the case of *boiling:freezing*, it is still expected that they will co-occur more frequently than chance predicts.

**Table 3.1** *Pairs along the* TEMPERATURE *continuum*[4]

| Word 1 | Word 2 | GOE-rating | T-Score | EAT (1/2) | EAT (2/1) |
|---|---|---|---|---|---|
| hot | cold | 1.13 | 20.32 | 0.64 | 0.34 |
| warm | cool | 1.84 | 5.78 | 0.02 | 0.09 |
| freezing | boiling | 1.975 | 2.22 | 0.01 | 0.01 |
| hot | cool | 2.68 | 5.22 | 0.04 | 0.22 |
| chilly | warm | 2.265 | 1.34 | n/a | 0 |
| cold | mild | 4.815 | 1.31 | 0 | 0.02 |
| lukewarm | tepid | 6.215 | 0.99 | n/a | n/a |
| chilly | steaming | 3.565 | 0 | n/a | n/a |
| sweltering | nippy | 3.275 | 0 | n/a | 0 |
| cold | friendly | 4.015 | -0.84 | 0 | n/a |
| wintry | sultry | 4.565 | 0 | n/a | n/a |
| scorching | glacial | 3.05 | 0 | n/a | 0 |
| bland | hot | 4.855 | 0 | 0 | 0 |

*Warm:cool* are much less restricted in their application and both lexemes have several sub-senses which do not contrast directly with each other, for example *warm* 'friendly' and *cool* 'stylish'. However, this pair's t-score is still relatively high and one elicits the other to some extent in both directions. Interestingly from an antonymy perspective, the EAT scores for *warm* show that it elicits *cold* much more frequently (0.28) than its 'direct' antonym; this also holds true for *cool* and *hot* (0.20). This is not altogether surprising since all four lexemes share a significant amount of semantic range and both *cold/cool* and *hot/warm* can be used interchangeably in many contexts. However, from a strictly theoretical perspective, the fact that these pairings are not equidistant from the midpoint of the scale should result in a significantly lower associative score. It seems that this only comes into play when the task is more focussed on lexical opposition as opposed to being a free association task. Despite the differing t-scores, with *freezing:boiling* barely displaying a significant co-occurrence pattern (2.22), the GOE-scores of the two pairs are remarkably similar. *Freezing:boiling* (1.97) is only slightly lower than *warm:cool* (1.84) and both score significantly worse than the base pair *hot:cold*.

A closer look at the semantic range of *boiling* and *freezing* (cf. Figure 3.3) might provide an insight into the low t-score. It seems that, despite their closely matching opposed sub-senses, the distribution patterns of the two lexemes differ considerably. Out of the three sub-senses in Figure 3.3 below, *freezing* mainly appears in sense 2 ('(very) cold') whereas sense 1a of *boiling* ('bringing/coming to

[4] If the order in which the pairs are presented in a table is not explicitly stated, they are ordered by assumed canonicity with the most canonical pairs first, then other symmetrical pairs and then asymmetrical and specifically constructed pairs. This does not always correspond to the GOE-rating scores but indicates roughly the expected ranking of the pairs under discussion.

the boil') is by far the most common one.[5] However, the LDCE and the OALD list *boiling* (adj.) as a separate sense which is overtly marked as the opposite of *freezing* in both dictionaries.

FREEZING
1a. to reduce temperature to achieve a change of state (liquid to solid)
*The **freezing** and canning process reduces the number of nutrients in food.* (BNC EX5 2006)
1b. (metaphorically) to halt something
*Between 1980/81 and 1982/83 the **freezing** of university places* […] (BNC FP4 479)
2. (very) cold but usually not literally at freezing point
[…] *doing one night shows for the troops in draughty halls and **freezing** aircraft hangars.* (BNC ACE 2923)

BOILING
1a. bringing something (mainly liquid) to the boil (transitive/intransitive)
*Return to heat and stir until **boiling**.* (BNC BN5 883)
1b. (metaphorically) uncontrollable emotions (anger/excitement)
*Gripping the rail, she struggled to keep a check on her **boiling** emotions.* (BNC HA6 472)
2. (very) hot
[…] *which was used for having tea in the shade on **boiling** afternoons.* (BNC AB4 1606)

**Figure 3.3** *Sub-senses of* freezing *and* boiling

This mis-match in the frequency of sub-senses reduces the number of occurrences in which the two lexemes can be found in close proximity and this results in a considerably lower t-score. The lexemes are, compared to *warm:cool*, also individually less frequent and this might bias the result of the t-score as this measure is more effective with a higher overall frequency.

However, overall, the results so far are in line with expectations if one assumes that criterion 1 (distance from the midpoint of a scale) plays an important role in antonym judgements.

The next group of lexemes to be discussed consists of five lexemes which would generally be called 'indirect' antonyms (cf. Charles et al. 1994); they contain one member of the base pair which has been matched with a synonym of the other member of the base pair. The pairs under discussion are *hot:cool, cold:mild, hot:bland, cold:friendly* and *warm:chilly*. It should be noted that despite the fact that not all of the five pairs include members of the base pair some of them, for example *warm:chill*y, nevertheless exhibit similar behaviour.

[5] Data taken from the BNC World (accessed November 2008 - www.natcorp.ox.ac.uk)

*Hot:cool* has the highest t-score of this group (5.22), which is roughly the same as for *warm:cool*. However, the GOE-rating (2.68) is considerably lower than that of *warm:cool* and is also lower than that of *warm:chilly* (2.26) despite the fact that *warm:chilly* has an extremely low co-occurrence score (1.34). The difference in frequency of co-occurrence should, if it is indeed the sole deciding factor, lead to a higher GOE-rating score for *hot:cool*. However, the scores for both lexemes are in roughly the same region on the scale and this is in keeping with the assumption that, if one of the above criteria is violated, in this case equidistance from the midpoint of the scale, the result is a proportionately lower GOE-score. This holds despite the fact that the t-score for *warm:chilly* is negatively influenced by the lower individual frequency of *chilly*.

*Cold:mild*, with a similar t-score to *warm:chilly* (1.31), scores much lower on the GOE-rating (4.81). This is presumably due to the fact that *mild* is not strictly speaking a synonym of *warm* or *hot* and therefore is neither directly nor indirectly opposed to *cold*. There are contexts in which the two occur in opposition (cf. example 3.1 below) but the contextual scaffolding would have to be reasonably strong and without it they are not readily identified as antonyms.

(3.1) *Is it cold out? No mild.* (BNC KBH 918)

Another factor which presumably influences the low rating of this pair is the difference in semantic range of the two lexemes. Although there is overlap in the temperature sense, both *cold* and *mild* have other sub-senses which do not stand in any antonymic relation. It might be that the core component of any good antonym pair, minimal difference, is stretched too far in this case.

*Cold:friendly* and *hot:bland* are slightly different examples of indirect antonyms as both are not based on an antonymic relation along the temperature scale but on a scale which is salient for certain sub-senses: ‘friendliness’ and ‘spiciness’. This, as discussed above, results in a mismatch of non-propositional meaning and a lesser amount of shared semantic range which in turn results in fewer opportunities for co-occurrence. Both pairs score below 4 on the GOE-rating and the t-score for *cold:friendly* is negative as both lexemes are extremely frequent individually but the number of their co-occurrences is disproportionately low. They are quite clearly not considered very good examples of opposition when used uncontextualised. However, as in so many cases, both can easily be used contrastively in certain contexts.

The discrepancies between t-score and GOE-rating discussed in the sections above can be seen clearly in Figure 3.4. The more subtle differences between t-score and GOE-rating can more easily be seen in Table 3.1 above but the graphs provide a clear visual representation of the differences under discussion. The graph with the diamond markers shows a gradual increase in GOE-rating scores not quite matched by the corresponding decrease in t-score which would be expected were we to consider t-scores 'perfect' predictors of antonymic strength. However, as will be shown in further detail in the following sections, the strength of prediction of a high t-score on GOE-rating results is seen most clearly in gradable adjectival antonyms such as the ones above.

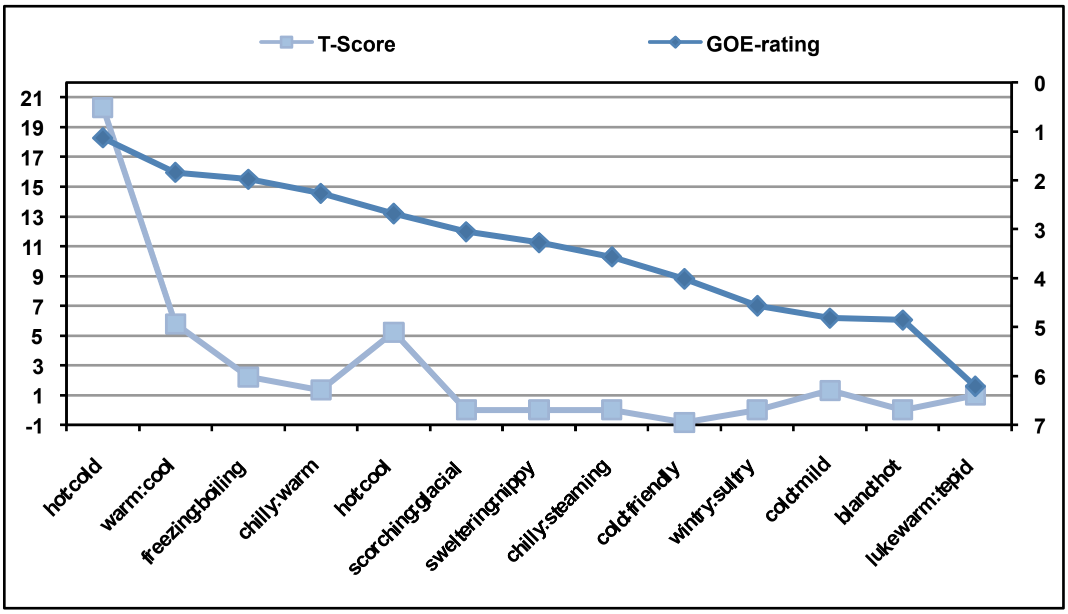



**Figure 3.4** *GOE-rating and t-score for pairs on the* TEMPERATURE *scale*

The last group in this section consists of another five pairs, three of which are considered roughly equidistant from the midpoint of the scale (*sultry:wintry, scorching:glacial* and *tepid:lukewarm*) and two are not (*sweltering:nippy, chilly:steaming*). Out of these pairs *scorching:glacial* scores highest whereas *tepid:lukewarm* scores extremely low. This is unsurprising, since *tepid* and *lukewarm* are usually used to describe the same range of temperature and are therefore frequently synonyms. *Scorching* and *glacial* display a similar degree of extremity in opposite directions which might account for the fact that this pair scores much higher (3.05) than the other symmetrical pair. They also have quite a large shared semantic range as they are opposed in several ways. *Wintry:sultry* (4.56), on the other hand, is a difficult pair of two lexemes which are individually not very frequent and definitely do not seem to have much shared semantic range despite the fact that they both describe weather conditions. They are rarely used in combination as the quality of the temperature or weather they describe is quite different. This might be a case of a difference in purity of opposition; *hot:cold* refer solely to temperature whereas *wintry:sultry* carry more features which might obscure the salient difference in temperature between the two conditions.

*Sweltering:nippy* and *chilly:steaming* both score around 3 in the GOE-rating and are therefore not considered good antonyms. However, they display enough antonym characteristics to be a recognisable pair of opposites. The BNC does not provide co-occurrence data for any of the lexemes in either of the pairs as they are quite infrequent. Neither of the pairs is symmetrical and there might also be a difference in purity of opposition, a factor that will be discussed in greater detail in the following two sections where it is more clearly visible due to the choice of test words.

Looking back on the discussion of the antonym pairs on the temperature scale, it has become clear that some of the criteria proposed by Cruse (1986), Muehleisen (1997) and Murphy & Andrews (1993) seem to be important factors in determining antonymic strength. Frequency of co-occurrence, and associative strength in general, clearly plays a role, but already in this very small sample of pairs there have been instances where the co-occurrence data cannot account for certain results in the GOE-rating. This will be investigated further throughout this chapter as well as in Chapter 4, where the results discussed here will be compared with the data collected from lexical decision tasks.

#### *3.4.1.2 The MERIT scale:* GOOD:BAD

The cluster of opposite pairs discussed in this section is based on the MERIT scale. The antonyms on this scale are considered overlapping antonyms as one member of the pair covers part of the scale of the other. This might play an important role when it comes to GOE-judgements of pairs which are non-canonical and asymmetrical. The discussion of the results will follow similar patterns to the one in 3.4.1.1 and will attempt to further consolidate the findings which resulted from the analysis of the pairs on the temperature scale.

Ten pairs are included in the GOE-rating questionnaire; four of these pairs are symmetrical, i.e. equidistant from the midpoint of the scale, four are asymmetrical and the remaining two, *good:disobedient* and *gracious:mean*, are constructed as ‘indirect’ antonyms of separate sub-senses of *good:bad*. The OALD does not mark any antonyms in its entry for *good* whereas the LDCE gives several, according to the different sub-senses (*bad*, *poor* and *naughty* but not *evil*). For *bad*, the LDCE gives *good* and the OALD, once again, does not explicitly refer to the antonymic relation between *good* and *bad*.[6]

**Table 3.2** *Pairs along the MERIT scale*

| Word 1 | Word 2 | GOE-rating | T-Score | EAT (1/2) | EAT (2/1) |
|---|---|---|---|---|---|
| good | bad | 1.175 | 27.19 | 0.80 | 0.56 |
| good | evil | 1.39 | 15.77 | 0.02 | 0.17 |
| excellent | atrocious | 2.14 | 0 | 0 | n/a |
| excellent | bad | 2.64 | -0.47 | 0.03 | 0 |
| mediocre | brilliant | 3.61 | 0 | n/a | 0 |
| good | disobedient | 3.375 | 0 | 0 | n/a |
| bad | satisfactory | 4.415 | 0 | 0 | n/a |
| fair | poor | 4.33 | 0.96 | 0 | 0 |
| good | mediocre | 4.815 | 2.59 | 0 | n/a |
| gracious | mean | 3.44 | 0 | n/a | 0 |

Table 3.2 shows that, as expected, *good:bad* scores very high in both GOE-ratings and frequency of co-occurrence scores. Interestingly, there is a noticeable discrepancy between the EAT scores for *good:bad* (0.80) and those for *bad:good* (0.56). Both are extremely high when compared to the results for other highly canonical antonym pairs but the evaluatively positive member of the pair (*good*) elicits *bad* with a frequency that is not repeated in any other pair. This corresponds to the preferred sequence of this particular pair and is in keeping with the criteria for sequencing proposed by Jones (2002).[7] The t-score of this pair is also one of the highest in the study; this is presumably influenced, in part, by the high individual frequency of the two members of the pair as well as by the great amount of shared semantic range. Both lexemes are applicable in a large number of contexts and thus their connection is very deeply entrenched.

*Good:evil*, which is usually also cited as a canonical antonym pair, receives slightly lower scores on both measures but nevertheless scores very highly. The slight drop in the results of the GOE-rating can be attributed to a smaller shared semantic range as the ranges covered by *good* and *evil* respectively differ markedly. *Evil* is only an antonym to one of the sub-senses of *good* ('morally acceptable') and, theoretically, this combination should therefore result in similar scores as other pairings which solely focus on one sub-sense (e.g. *good:disobedient, good:naughty*). However, the fact that, in certain contexts, this pair is highly canonical (also as a noun pair) ensures that the associative strength is sufficient to balance this difference in semantic range.

[6] Antonym marking in learner dictionaries is interestingly unsystematic and has been investigated by Paradis & Willners (2007) (Collins COBULID) and myself (OALD).
[7] For a list of criteria which influence the sequencing of antonyms, see Jones (2002), 2.2.2 and 6.1.2.3.

The two other symmetrical pairs, *fair:poor* and *excellent:atrocious*, score very differently in the GOE-rating. While *excellent:atrocious* falls into group 2 (good opposite), *fair:poor* receives a ranking of below 4, which is extremely low. In terms of its distance from the midpoint of the scale as well as its GOE-rating score, this pair is reminiscent of *lukewarm:tepid* in 3.2.1.1 above. However, in this case the two lexemes do not denote the same portion of the scale but apply to neighbouring stretches. Very often, when there is only a tripartite distinction (e.g. *good – fair – poor*), *fair* and *poor* cannot be considered equidistant from the midpoint of the scale. In some cases, where a fourth adjective is added to the bottom of the scale (*good – fair – poor – terrible*), the midpoint would move in between the two lexemes. Despite this possibility, it seems that the opposing relation lacks the necessary salience to make the pair a good example of an antonym pair. A large number of occurrences found by Google seem to contain *fair* and *poor* in a three-part judgement scale together with *good*. Furthermore, the fact that the overall distance between the two concepts is lower than in other examples in this section might further contribute to the low GOE-rating score. Both lexemes also have other readings which are completely unrelated and, in both cases, are the lexemes' primary meanings (*fair:unfair* and *poor:rich*).

*Excellent:atrocious*, however, is rated as a good antonym pair in the GOE-rating (2.14) despite the fact that no co-occurrence data could be obtained from the BNC. Both lexemes mark the extremes of the scale and, as a consequence of this, they lose the gradability which is central to the category of antonym (in the sense of Lyons 1963). Gradable modification through lexemes such as *very, quite, extremely, and reasonably* cannot be used with words which denote the extremes of a scale (see Table 3.3).

**Table 3.3** *Strong and weak adjectives*

| | very | quite | extremely | absolutely |
|---|---|---|---|---|
| good | Y | Y | Y | **N** |
| bad | Y | Y | Y | **N** |
| excellent | **N** | **N** | **N** | Y |
| atrocious | **N** | **N** | **N** | Y |
| hot | Y | Y | Y | **N** |
| cold | Y | Y | Y | **N** |
| freezing | **N** | **N** | **N** | Y |
| boiling | **N** | **N** | **N** | Y |

The reverse is also true as *absolutely*, for example, cannot be used with base adjectives (e.g. *good*) as it possesses an inherent quality of absoluteness which is not present in any base adjective. The question whether this loss of gradability plays a role in the judgement of this pair (and other similar ones like, for example, *boiling:freezing* and *huge:tiny*) cannot be answered satisfactorily at this stage as the fact that all these pairs have significantly lower t-scores could equally well be the reason for the lower GOE-rating. However, the influence of gradability on antonym judgements will be explored further in Chapter 6.

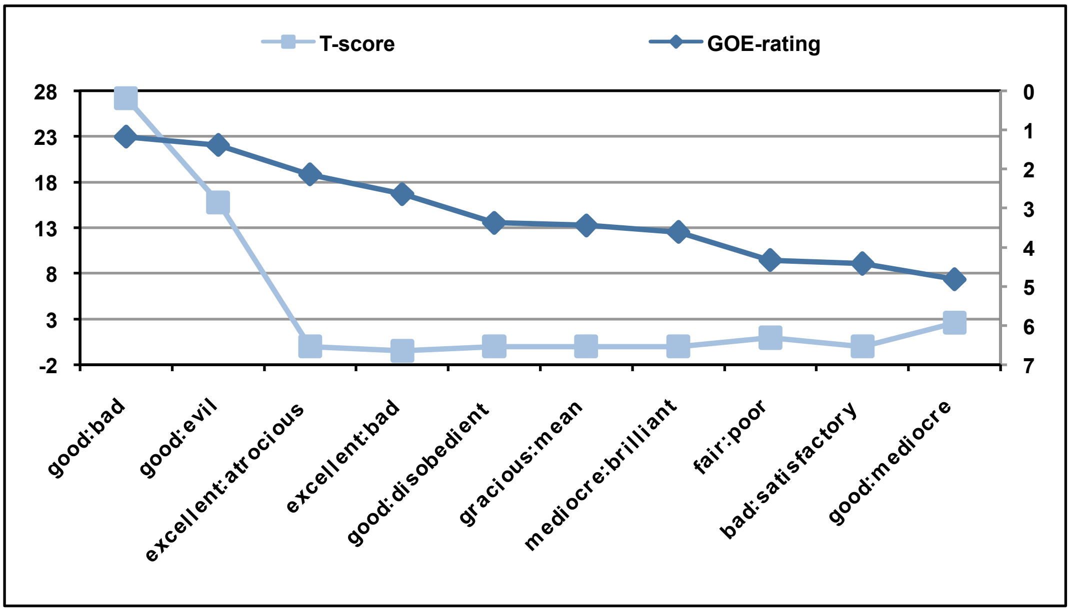



**Figure 3.5** *GOE-rating and t-score for pairs on the* MERIT *scale*

On the MERIT scale, the graph plotting GOE-rating versus t-scores (cf. Figure 3.5) shows a greater correlation between the two measures than in most other cases. Nevertheless, despite the fact that the GOE-ratings in the latter part of the graph still increase consistently, the graph representing the t-score seems to have levelled off. This could, however, be explained by stating that below a certain t-score, the co-occurrence is too minimal to contribute significantly to antonymic strength. Overall, it looks as if the adjectival pairs on the MERIT scale are more reliant on frequency of co-occurrence than the other pairs on the TEMPERATURE scale discussed in the previous section.

As for the 'indirect' antonyms, the first pairs to be considered are those which still contain one member of the base pair: *good:mediocre*, *excellent:bad* and *bad:satisfactory*. Out of these pairs, *good:mediocre* is the only one which has a significant t-score, as *excellent:bad* in fact scores negatively, presumably due to the high individual frequency of both lexical items. However, both *excellent:bad* (2.64) and *bad:satisfactory* (4.41) score higher than *good:mediocre* (4.81) in the GOE-rating. One possible explanation for the higher score of *excellent:bad* is the greater distance between those lexemes on the merit scale. In both of the other indirect pairs, the base lexeme is paired with a lexeme which expresses mediocrity (*mediocre, satisfactory*) and thus does not have the same amount of conceptual distance from the respective base lexeme, which in turn might reduce the salience of the difference between the two words/concepts (cf. *fair:poor*, *mild:cold* and *warm:chilly*). This would have to be validated by looking at other similar pairings to ascertain whether there are other possible factors, for example a difference in raw individual frequency (*excellent*: 6620; *mediocre*: 174; *satisfactory*: 2161)[8] which would, in this case, also account for the three different GOE-rating scores.

None of the last three pairs has a t-score as there were no instances in which they co-occurred in the BNC. Despite this fact, all three score in the medium range on the GOE-rating questionnaire (3.37-3.61). *Mediocre:brilliant* is not as asymmetrical as one might assume and there are a number of instances on Google where the two lexemes are used together in the same sentential contexts as most other antonym pairs (cf. Example 3.2), unlike *poor:fair,* which mainly occurred in rankings together with *good* and/or *excellent*.

(3.2) *I feel anyone can design (whether* mediocre *or* brilliant*).*[9]

This co-occurrence pattern seems to be a factor contributing to the comparatively high GOE-rating score despite both low individual frequencies and low frequency of co-occurrence.

The remaining two pairs are those mentioned at the beginning of this section. They are constructed using one sub-sense of *good* in the case of *good:disobedient,* but rather than taking the most common opposing lexeme (*naughty*) choosing one which is a little further removed and differs in style as well as morphological makeup (polymorphemic containing a negative prefix). This seems to influence the judgement of goodness-of-exemplar rating as *disobedient* is not the primary partner for *good* and, in addition to that, has its own antonym: *obedient*. However, the rating is still within 'antonymous' range as the conceptual opposition is salient enough, especially in certain contexts (e.g. *a good child* vs. *a disobedient child*). Like some cases in 3.2.1.1, this pair needs some degree of contextual scaffolding to become a fully-fledged antonym pair as the connection is otherwise too oblique.

*Gracious:mean* is quite far removed from the original pair as it is constructed from fairly distant synonyms of both members of the base pair. It was slightly surprising to see it rank reasonably high in the GOE-rating. The two lexemes, despite not being canonical antonyms by any means, seem to incorporate some

[8] taken from the BNC
[9] http://fashion.meetup.com/313/calendar/8071968/ (accessed December 2008)

opposition which must be contextually supported. Judging by their relatedness to the base pair and a co-occurrence score of zero in the BNC as well as a lack of associative strength between the lexemes in the EAT data, this pair should have been judged considerably lower. The data from the lexical decision task to be discussed in the following chapter will show whether this high ranking stems from an attempt to rationalise and make sense of the pair or whether it really has more antonymic character than expected.

Overall, the results of the analysis of this cluster of antonym pairs support the conclusions drawn in the previous section: frequency of co-occurrence is, without a doubt, an important factor in the judgement of antonyms. However, there are other factors which seem to play a role, for example the distance from the midpoint of a scale, and indeed the conceptual distance of the lexemes from each other. Semantic range can also still be considered an important factor despite the fact that it seemed less central in the discussion of *good:bad*. This raises the question whether there is a difference in the judgement of evaluatively subjective antonym pairs as opposed to objective pairs (*hot:cold, long:short, big: small*).

#### *3.4.1.3 Matters of size:* BIG:LITTLE *and* SMALL:LARGE

The last cluster of gradable antonym pairs which will be discussed in this chapter is that around the base pairs *big:little* and *small:large*. Before looking at the data in detail, the curious case of the relationship between the base pairs above will be examined a little more closely.

English finds itself in a somewhat special situation where the size continuum is concerned, in that it has two antonym pairs of nearly equal standing which can be used interchangeably in some contexts (with very little connotational difference). The two lexemes which denote the upper end of the scale (*large, big*) are often used interchangeably and are generally considered to be very good synonyms. The same applies to *small* and *little* at the opposite end of the size continuum. However, native speakers will commonly state that the antonymic combinations in which the four lexemes appear are relatively fixed (cf. Figure 3.6).

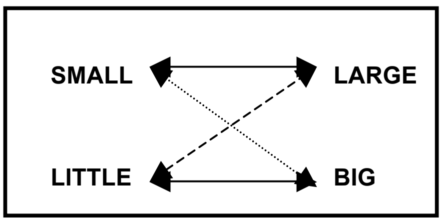



**Figure 3.6** *Antonymic relationships between* big, large, small *and* little

Figure 3.6 illustrates the combinations of lexemes which commonly occur. The two conventional pairs (*big:little* and *large:small*) are connected with unbroken arrows whereas the dotted line for *big:small* indicates that the combination is judged better than that between *large:little* (dashed line) but is usually not considered one of the conventional pairings.[10] Native speaker intuitions seem to be supported by empirical evidence. K. J. Miller (1998: 52), for example, states that '[…] the pair *large* and *little* are simply not accepted as antonyms. Overwhelmingly, association data and co-occurrence data indicate that *big* and *little* are considered a pair and *large* and *small* are considered a pair.'

Several researchers working on antonymy have commented on the phenomenon illustrated above. Gross et al. (1988) and Miller (1998) state that this, in their view, gives support to the theory that antonymy is an associative relation which exists only between certain lexical items rather than between the concepts attached to those items. The claim that '*large/little* contrast conceptually just as sharply as do [...] *large/small*, and yet they are not direct antonyms' (Gross et al. 1988: 3) is used to equate associative strength with antonymic strength.

Muehleisen 1997, however, investigates these pairs more closely and conducts a thorough analysis of their semantic range, which serves to illustrate their distributional differences.[11] In the discussion of the semantic range analysis, Muehleisen (1997: 104ff.) shows that *little* does not share a large amount of semantic range with *large*, whereas *small* seems to occur in similar contexts and collocational combinations much more frequently. The case Muehleisen makes for the strong reciprocal connection between *big:little* is not as convincingly supported by the data as in the case of *large:small*. *Big* also occurs in numerous contexts in which *little* would not be used, but Muehleisen claims that despite the fact that '[…] from the point of view of *big*, *little* is not a particularly good candidate as an antonym, from the point of view of *little*, *big* is a very good candidate.' (1997: 108ff.) She supports this by saying that *big* is the better match for *little* as these two lexemes share more semantic range than *large:little* and the fact that the connection between *large:small* is such a strong one, which leaves *big* as the only available option for *little*. This seems a little more tenuous than the explanation for the first pair of the size continuum.

Furthermore, Muehleisen uses corpus data to show that, statistically speaking,

[10] This section is based on discussions with the participants of the questionnaire study as well as conversations with other native speakers of English who were asked to pair up the four lexemes much as I have done in Figure 3.3.
[11] For a detailed analysis of the four central lexemes on the size continuum, see Muehleisen (1997, Chapter 2 - http://www.f.waseda.jp/vicky/dissertation/pdf.html).

*big* and *small* have more semantic range in common than *big* and *little* (1997: 112) and that therefore, *big:small* should be the better antonym pair. As an explanation for the seemingly stronger connection between *big:little*, she uses reciprocity, as *small* is found primarily in the same environments as *large* and shares less semantic range with *big* but still more than *little* shares with *big*. Whether the antonymic relationship between *big:little* is indeed perceived as stronger than that between *big:small* by the participants in the GOE-rating task will be discussed below.

Another informative source of conventionalized usages is provided by learners' dictionaries as, due to their purpose, there is usually a particular focus on providing the most essential definitions and contrasts between usages. Muehleisen also considered dictionary entries in her analysis and they often show considerable variation in which lexemes they mark as having an opposite. *Big*, in the OALD (7th edition, 2005), is not overtly marked as having any opposite at all while a previous version of the OALD describes *small* as the usual opposite of *big* and *large* (cf. Muehleisen 1997: 69). The 7th edition discusses the difference between the synonyms *great, large* and *big* regarding their collocational distribution but without reference to their antonyms. None of the other three lexemes in the base pairs is overtly marked in the OALD. In the LDCE (2003), *large* has *small* marked as an opposite and while, much like in the OALD, some of the lexemes are defined in terms of the other (e.g. *small* – 'not large'), no other overt markings are indicated.

The next section will deal with the closer analysis of the GOE-rating data for eleven pairs along the SIZE continuum (Table 3.4); ten of these pairs are adjectives and one, *giant:dwarf,* is a noun pair. Once again, they were chosen to deliberately violate some of the criteria for good opposition and, in this case, all possible combinations of the central four lexemes were included to investigate their complex relationship.

**Table 3.4** *Pairs along the SIZE continuum*

| Word 1 | Word 2 | GOE-rating | T-Score | EAT (1/2) | EAT (2/1) |
|---|---|---|---|---|---|
| big | little | 1.24 | 8.13 | 0.18 | 0.20 |
| small | large | 1.2 | 25.67 | 0.26 | 0.45 |
| big | tiny | 2.04 | 1.82 | 0 | 0.02 |
| small | huge | 2.25 | 1.53 | 0.01 | 0 |
| small | tall | 3.14 | 2.55 | 0.05 | 0.06 |
| giant | dwarf | 1.45 | 1.72 | 0.08 | 0.13 |
| large | little | 2.05 | -3.72 | 0 | 0.01 |
| little | gigantic | 2.19 | 0.42 | 0 | n/a |
| big | small | 1.225 | 10.79 | 0.29 | 0.20 |
| tiny | huge | 1.49 | 3.90 | 0 | 0 |
| colossal | miniscule | 1.505 | 0 | n/a | n/a |

After looking at the symmetrical pairs, the analysis will move to the asymmetrical and more peripheral pairs. However, in the case of the size continuum, the focus of the analysis is slightly different from the previous two cases as the centre of the category is more thoroughly investigated. Symmetry, as a characteristic of good antonyms, has already been considered in the analysis of the TEMPERATURE and MERIT continua and will therefore only play a minor role in this section.

Having discussed the conventional combinations of *big*, *large*, *small* and *little*, I would now like to compare previous ideas and findings to the data presented in this chapter. The pairs which will be analysed in this section are: *big:little*, *small:large*, *big:small* and *large:little.* While *small:large* (25.67) has by far the highest t-score (compared to 8.13 and 10.79 for the second and third pair above), the GOE-rating scores for the first three pairs are almost identical (1.24, 1.2 and 1.22 respectively). There seems to be no difference in goodness of antonym judgement between *big:little* and *big:small*, the former being the conventional pair and the latter, according to Muehleisen (1997), displaying the better match of semantic range. The EAT scores show that *big* elicits *small* (0.29) much more often than it elicits *little* (0.18) and that both *little* and *small* elicit *big* the same number of times (0.20). *Small*, on the other hand, elicits *large* (0.26) more often than *big* (0.20) and *little* does not elicit *large* at all.

The data discussed above seems to indicate that there are three central pairs on the SIZE continuum, as *little:large,* with a GOE-rating score of 2.05 and a very low t-score which does not indicate any significant co-occurrence, does not follow the same pattern. The fact that there is no significant difference between *big:small* and *big:little* is interesting since the conventional pairings are still readily identified by native speakers. Highly proficient non-native speakers, however, will often prefer *big:small* as a better antonym pair.[12] The large overlap in semantic range and, as a consequence of this, the high frequency of co-occurrence seems to determine this pattern as *little:large* simply does not seem to have enough shared range to allow the lexemes to co-occur. The results of the lexical decision tasks discussed in Chapter 4 will show this pattern more clearly (cf. 4.2.4).

This is also reflected in the visual representation of GOE-rating data and t-scores in Figure 3.7. The pattern of t-scores shows a large amount of variability even when GOE-ratings are fairly stable. This differs considerably from the distribution seen in Figure 3.5 for the MERIT scale and from the graphs for the TEMPERATURE scale (Figure 3.4). It seems that beyond a certain t-score, the impact

[12] A small survey among highly proficient speakers of English showed that most considered *big:small* an excellent opposite pair. These informants paired *small* with *big* as well as with *large*.

of an increase in frequency of co-occurrence does not directly result in an increase of antonymic strength since, as already discussed above, the central pairs show almost identical results in the GOE-rating but one pair, *small:large*, has a significantly larger co-occurrence ratio. This, taken together with the data from the MERIT scale, seems to suggest that GOE-ratings are only susceptible to changes in t-score within a certain range whereas differences in t-scores which are either extremely low or extremely high do not produce correspondingly large changes in GOE-rating.

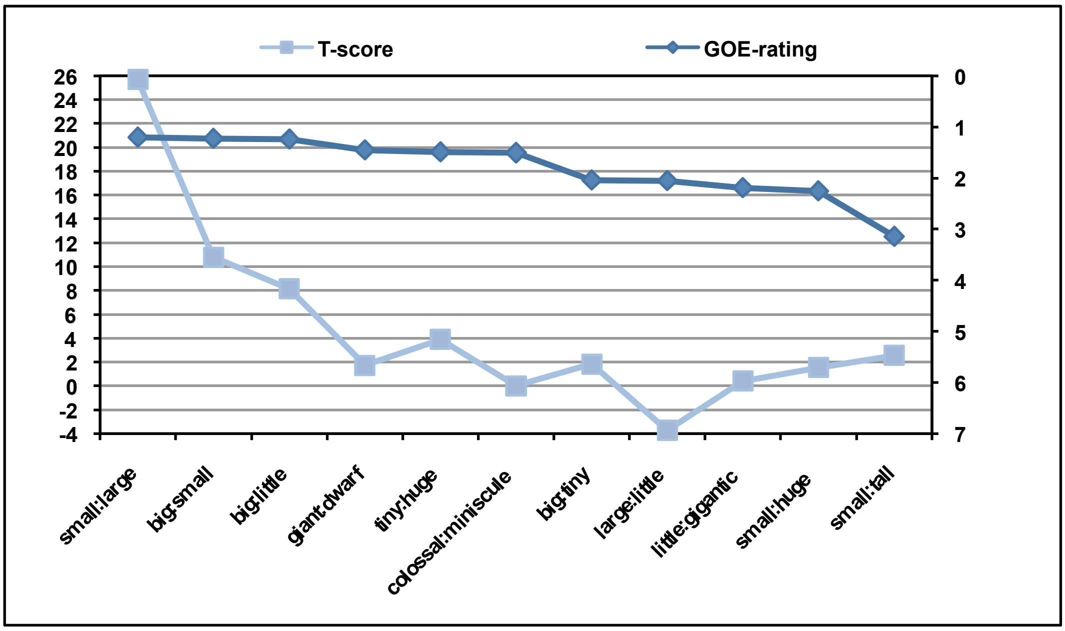



**Figure 3.7** *GOE-rating and t-scores for pairs on the SIZE continuum*

*Huge:tiny* and *colossal:miniscule* are the two non-central pairs which are nevertheless symmetrical. Both score relatively high on the GOE-rating (at 1.49 and 1.50 respectively), especially considering the scores of similar pairs from the previous two clusters of opposites (*freezing:boiling* 1.97, *warm:cool* 1.84, *excellent:atrocious* 2.14), despite the fact that neither have very high t-scores. In fact, no co-occurrence of *miniscule:colossal* is attested in the BNC and neither word is in the EAT. The EAT scores for *tiny:huge* are zero despite both lexemes being in the EAT. Judging by the GOE-rating scores, both pairs seem very robust opposites, as good as some which are generally called canonical, and yet they do not display some of the features which are deemed necessary for canonical antonymy.

Before moving on to the last group of items, the three asymmetrical pairs in the questionnaire - *big:tiny*, *small:huge* and *little:gigantic* – will be analysed. All three contain one member of the base pairs and one member located further towards the extremes on the size continuum. All three of these pairs score just above 2.0 on the

GOE-rating (2.04, 2.25 and 2.19 respectively) with very low t-scores which do not indicate any significant co-occurrence beyond that predicted by chance (1.82, 1.53 and 0.42). Many pairs which achieve similarly low t-scores, morphologically related pairs aside, score much lower on the GOE-rating despite conceptual opposition (cf. 3.4.1.1 *chilly:warm*, *cold:mild*). The only 'structural' difference between these pairs is that the members of the asymmetrical pairs on the size continuum are a greater distance from each other than the pairs on the temperature scale since they consist of the pair of a base member and one that is closer to the centre of the temperature scale. Overall distance between the lexemes, and therefore concepts, seems to be a factor in the judgment of asymmetrical antonym pairs as this, evidently, increases the contrast between the two concepts. This will be further investigated in 4.3.1.4 and 4.3.3.[13]

To conclude this section on antonym pairs on the size continuum, two pairs will be considered which are different from the ones analysed above. One is a nominal pair, *giant:dwarf*, and the other is an indirect pair in which the second member does not, strictly speaking, belong to the domain of SIZE but to that of HEIGHT. In German, the same lexemes are used for both SIZE and HEIGHT (when describing people). English, however, uses distinct lexical items for those two concepts. *Small:tall*, even though both lexemes can be used in certain common contexts and in collocation with the same lexical items, do not make a very good antonym pair. The GOE-rating score is 3.14 and the t-score is 2.55, indicating a higher co-occurrence than in the cases of those discussed in the previous section, but nevertheless not one much higher than chance would predict. The EAT scores show that they do elicit each other relatively frequently (at 0.05 and 0.06), compared to other pairs which score much higher on the GOE-rating, but neither is the first association of the other. Even though in some contexts they are conceptually opposed, these are not the contexts they are primarily used in and both lexemes have one, or in the case of *small* even two, lexemes that they stand in primary opposition to.

*Giant:dwarf*, the only nominal pair discussed in this section, scores surprisingly high on the GOE-rating (1.45) despite the fact that the criterion of purity of opposition does not hold. The pair contains other features apart from the salient distinction *big:small*; a fact which should lead to a weakening of the salient axis. However, in this case the pair, despite a low t-score, is not judged much lower than

[13] As the temperature scale stimuli do not contain any combinations of lexemes structurally comparable to those on the size continuum, additional pairs will be included in the lexical decision task and GOE-judgments will be obtained for those pairs.

the central pairs of this domain. One reason for this could be that *giant:dwarf* incorporate other oppositions apart from *big:small*, for example *strong:weak*, which strengthen the opposition of the two concepts rather than incorporating other additional information which would lead to a weakening (cf. 3.4.3). This hypothesis will be investigated further in 3.4.3 when *male:female* is under discussion.

Summing up the conclusions drawn from the data on gradable antonym pairs, it seems that there are several factors, beyond associative strength, which play a role in the judgement of goodness of antonymy. By investigating cases in which a discrepancy between the different types of data could be seen, the following criteria (in no particular order) seem to be influential:

(1) symmetry
(2) distance between concepts on the continuum
(3) purity of opposition
(4) frequency of co-occurrence
(5) semantic range
(6) morphological relatedness

These criteria are, for the most part, those which have been mentioned before. Most of them are paradigmatic criteria but (4) and (5) are syntagmatic criteria for what is essentially a paradigmatic lexical relation. This extends the analysis further, especially when more peripheral lexemes are taken into account. As this section has largely covered ground which has been examined before, the fact that the criteria seem to be similar is not surprising. The data analysis of converses and *male:female* in the following two sections, however, promises to extend the list of factors influencing antonymic strength. Whether these additional criteria are also of importance when judging gradable (canonical) pairs remains to be investigated.

### 3.4.2 Converses

It has already been mentioned above that converses are, in Cruse's and Lyons' classifications, a sub-category of directional opposition. The directionality is always retained (cf. Cruse 1986) but not always literally; all converses nevertheless have a directional component. Despite their inclusion in his classification of opposition, Cruse is not certain that they are truly real opposites:

> While there is little doubt that among converses [...] are to be found several intuitively satisfactory opposites (*above:below, husband:wife* and *buy:sell*, for instance), there are also many that either, like *gift:donor*, are rather feebly opposed, or are not opposites at all, like *charge:cost*. The question must be raised, therefore, of whether direct converseness or indirect converseness is in itself to be considered a genuine type of oppositeness. (Cruse 1986: 239)

In the above quote, Cruse cites the fact that some converses are bad examples of opposition as a reason to question whether the notion of converseness can justifiably be included in a classification of opposition. However, some converses behave in very similar ways to other types of opposition and there are, after all, also better and worse examples of other types of opposites. This section will introduce the notion of converses from a lexical semantic (logical) approach (Cruse 1986). The converse data from the GOE-rating questionnaire will then be analysed in two parts; nominal and verbal converses will be treated separately as their judgement results display different patterns.

#### CONVERSES AS OPPOSITES

Cruse 1986 distinguishes between three different syntactic categories of converses depending on the valency of the lexeme involved. The types of converses are illustrated in Figure 3.8 below. For the purpose of this study the difference between the three types is not of importance and the distinction is only introduced for completeness' sake.

Converses are unique among antonyms; each member of the pair describes the same situation from a different perspective (or direction) and the concepts encoded by these lexemes are logically equivalent. As already mentioned above, the directional aspect of converses is a key feature which is always present. However, converseness is, like some other antonym types we have encountered, not a strictly logical relation. To use one of Cruse's examples (Cruse 2000: 172), *victim* can have several logically possible converse partners, e.g. *murderer*, *rapist, mugger* or *burglar* to name but a few. There are, nevertheless, also examples of converses which only have one possible partner; *husband:wife* and *above:below* are examples of this. Furthermore, the asymmetrical logical relation necessary for many converses excludes intransitive verbs and most adjectives as possible candidates for converseness (Cruse 1986: 232). Therefore, the only possible lexical categories for converses are nouns and transitive verbs and some prepositions.

| | | |
|---|---|---|
| **two-place converse** | A<br>B | *A is **above** B.*<br>*B is **below** A.* |
| **three-place converse** | | *Sam **lent** the car to Jack.*<br>*Jack **borrowed** the car from Sam.* |
| **four-place converse** | | *Sam **sold** the book to Jack for £1.*<br>*Jack **bought** the book from Sam for £1.* |

**Figure 3.8** *Types of converses*

The distinction between two, three and four place converses does not seem to make a difference in goodness of antonymy judgements and will therefore not be discussed in the analysis. The variation in logical reciprocity will, however, be examined in greater detail as this might have some bearing on the amount of antonymic strength of the converses in question.

One striking observation in the GOE-rating of all converses, verbal and nominal alike, is the overall greater standard deviation in the results for each pair as well as the greater discrepancy between the two possible sequences of each pair.[14] The latter difference does not, as far as this can be established, stem from a greater preference for a certain sequence but simply from a greater amount of indecision on the part of the participants as to whether, for each informant, converses are part of antonymy. They do seem to have a particular standing within a speaker's awareness and therefore result in judgments different from those of 'regular' opposites. The tables which show an overview of the data therefore include separate measures for both sequences of each pair as well as the mean of the two scores.

#### *3.4.2.1 Verbal converses*

This section will look solely at the verbal converses included in the GOE-rating questionnaire. Table 3.5 provides an overview of the scores for these pairs. Out of the ten converse pairs, five are generally considered good converses whereas scores for the other five vary considerably. The first group, which forms the core of a category of verbal converses, will be discussed first before less canonical cases are considered.

According to the informants who completed the GOE-rating, the best converse pairs are *defend:attack*, *give:take*, *buy:sell* and *come:go*. All these pairs score better than 1.8 which puts them in group I, the category of excellent opposites. However, none of them score as highly as some of the other canonical antonyms which have been discussed so far, despite the fact that the t-scores of *come:go* and *buy:sell* are extremely high. Interestingly, the converse pair which scored best overall (*defend:attack*) has a very low t-score (3.07) and medium associative strength in the EAT results. Sequencing seems to play a stronger part in some cases (*come:go*) than in others where there is virtually no difference between the two orders of presentation.

[14] In the questionnaire data overall, the standard deviation is relatively low and decisions seem largely homogeneous. However, with items which are rated in the middle of the continuum there is a larger amount of indecision and thus difference in the scores. This tendency seems to be particularly strong in the case of converses.

**Table 3.5** *Verbal converses (by overall GOE-rating)*

| Word 1 | Word 2 | GOE 1/2 | GOE 2/1 | GOE Overall | T-Score | EAT (1/2) | EAT (2/1) |
|---|---|---|---|---|---|---|---|
| defend | attack | 1.50 | 1.38 | 1.44 | 3.07 | 0.16 | 0.12 |
| give | take | 1.45 | 1.53 | 1.49 | 5.356 | 0.37 | 0.21 |
| sell | buy | 1.43 | 1.92 | 1.67 | 13.99 | 0.54 | 0.42 |
| come | go | 1.65 | 1.92 | 1.785 | 15.74 | 0.39 | 0.09 |
| borrow | lend | 1.98 | 2.92 | 2.45 | 4.67 | 0.52 | 0.43 |
| chase | flee | 3.33 | 3.59 | 3.46 | 0 | 0 | 0 |
| steal | donate | 4.10 | 4.18 | 4.14 | 0 | 0 | 0 |
| hunt | escape | 4.80 | 4.73 | 4.765 | 0 | 0 | 0 |
| purchase | trade | 5.48 | 5.13 | 5.305 | 0.985 | 0 | 0 |
| let | rent | 5.93 | 5.30 | 5.615 | 3.26 | 0.04 | 0 |

*Lend:borrow* was a pair which, despite very high associative strength in the EAT results and a significant t-score, placed surprisingly low in the GOE rating. Despite a higher frequency than *defend:attack* and much stronger associative scores, it only scores at 2.45 and the scores for the two sequences show a difference of almost one whole point. Had this study been carried out in a dialect area in which *lend* and *borrow* are used interchangeably (e.g. certain areas of Wales) or where one of the two lexemes is used to cover both directions of lending/borrowing (as is mostly the case in German), the results above would not have been surprising in the least. However, most informants declared themselves speakers of Standard British English, General American and Canadian English and would usually use *lend:borrow* as a pair of converses.

*Chase:flee*, *hunt:escape* and *steal:donate* are pairs which are less commonly accepted as opposition or even good examples of converseness. They all score relatively low on the GOE-rating and there are no co-occurrences of any of the pairs attested in the BNC; their associative strength is low as none of the lexemes elicit each other in the free word association task on which the EAT is based. Considering the fact that all other scores determining associative strength are non-existent, the GOE-rating is still comparatively high.

The last two verbal converse pairs – *rent:let* and *purchase:trade* – were constructed much like the 'artificial' pairs in the previous section. *Rent:let*, the lowest scoring converse (including nominal converses) is similar to *tepid:lukewarm* as the two lexemes in the pair are in many contexts essentially synonymous. *Rent* and *let* can in some contexts even be considered antagonyms (lexemes with contrasting sub-senses).[15] Furthermore, both members of this pair cover, in part, the same

[15] A more detailed discussion of antagonyms can be found in Lutzeier (1997 & 2001) who provides a broader overview.

concept and ranges of meaning. This pair, however, does retain some of its opposition in certain contexts. Another factor which might influence the judgement of this pair is the particularly striking difference in semantic range. The lexeme *let* is one of the entries in the OED with the largest number of sub-senses, and therefore *rent* only covers a minute part of the semantic range of *let*.

*Purchase:trade*, on the other hand, is a pair constructed using synonyms of *buy:sell*. This does not seem to work well with converses as the directional aspect, which is an important part of the oppositeness encoded in converses, seems to be less important. Furthermore, *purchase* and *trade* are not only different in register but do not seem to be very well matched for semantic range. This example shows that some of the same criteria for good opposition, in this case semantic range (Muehleisen 1997) or match of non-propositional meaning (Cruse 1986), also play a role in determining a converse's place on the continuum of opposition.

Overall, verbal converses seem to follow a similar pattern to that followed by other opposites: distribution along a continuum from excellent opposition to poor (or non-) opposition. This seems to indicate that, as they behave similarly, they should be considered part of the larger category of opposition, especially as many of the sub-types in Cruse's classification do not seem to result in different judgements as to how readily certain pairs are identified as opposites.

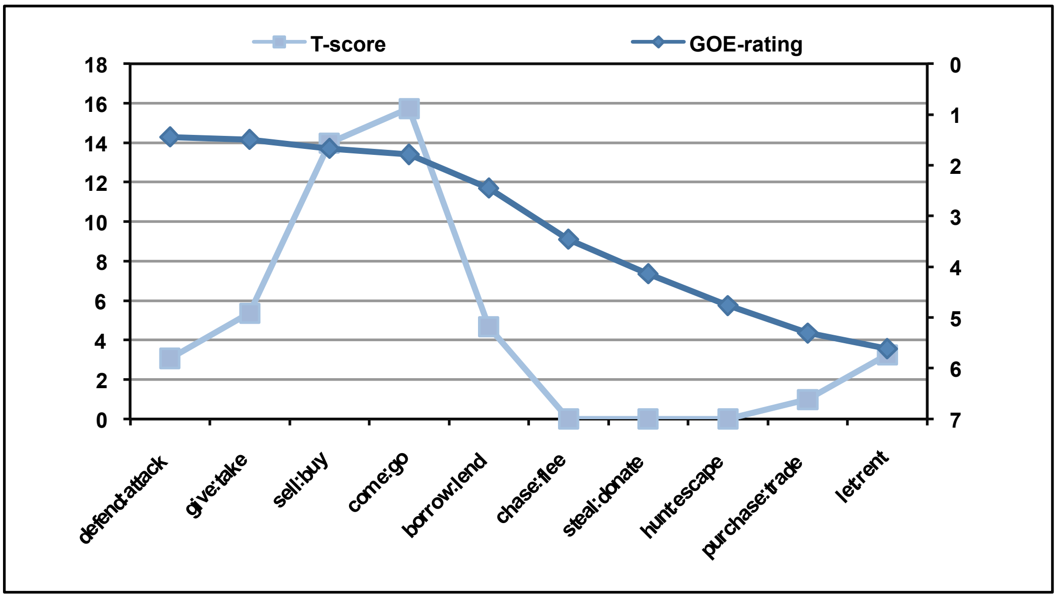



**Figure 3.9** *GOE-rating and t-scores for verbal converses*

Figure 3.9 shows that the graph representing the GOE-rating scores for verbal converses is very similar in shape to that of the gradable adjectival opposites showing a gradual increase and thus confirms the similarity in distribution mentioned above. The t-scores are less consistent in this case and do not seem to lend themselves to making accurate predictions about a verbal converse's place on the antonym continuum.

To determine how converses compare to other types of opposites with similar t-scores, Table 3.6 provides a list of the five highest-scoring converses and selected opposite pairs matched for t-score only. It is evident from the table below that the t-scores are not the strongest factor in determining judgements of converseness. The low t-score of *defend:attack* which has already been mentioned does not seem to be detrimental to the GOE-rating score of this pair while the matched pair, *dry:sweet*, scores significantly lower. *Dry:sweet* is a pair which needs to be provided with a very strong and specific context to be successful as an antonym pair and the two lexemes, when used to describe wine, are without a doubt excellent opposites. However, in a context-free task such as the GOE-task the other meanings of these words distract from this very specific shared semantic range. The small amount of shared semantic range also results in fewer co-occurrence possibilities. Whether high frequency of co-occurrence is a consequence of a large proportion of shared semantic range or a factor in its own right is a question which this and the following chapters endeavour to answer.

**Table 3.6** *Comparison of verbal converses and canonical antonyms*

| W1 | W2 | GOE | T-SCORE | W1 | W2 | GOE | T-SCORE |
|---|---|---|---|---|---|---|---|
| buy | sell | 1.67 | 13.99 | true | false | 1.01 | 13.52 |
| defend | attack | 1.44 | 3.07 | dry | sweet | 4.75 | 3.06 |
| give | take | 1.49 | 5.356 | little | big | 1.24 | 7.35 |
| come | go | 1.785 | 15.74 | old | young | 1.28 | 17.52 |
| lend | borrow | 2.45 | 4.67 | badly | well | 2.28 | 4.62 |

The opposite case is also shown in the table above; *buy:sell* and *true:false* have almost identical t-scores (13.52 and 13.99 respectively) but *true:false* is the pair which scores highest overall (1.01), whereas *buy:sell* is barely within the category of excellent antonymy. This comparison underlines a point which has already been made in previous sections. Frequency of co-occurrence, while without a doubt a factor contributing to associative strength, is not the factor which most strongly influences antonymic strength. This discussion will be continued in Chapter 4 where the results of the lexical decision tasks will provide more support for this argument and a regression analysis will afford more precise insight into what the deciding factors might be.

One hypothesis, which will be revisited after discussing nominal converses, is that word class is a factor in antonym judgement. It is possible that this is related to Cruse's (1986) criterion of purity of opposition. Adjectives are cognitively simplest and therefore purest.[16] They do not have elaborate cognitive models associated with them and therefore the opposition remains the most salient aspect. On the other hand, where verbs and nouns are concerned the associated models contain a larger amount of additional, extra-linguistic knowledge which could reduce the salience of the opposition. This possibility will also be investigated further in 3.4.3 where the antonym cluster around *male:female* is the object of discussion.

*3.4.2.2 Nominal converses*

As already mentioned above, the decision to discuss nominal and verbal converses separately is based on the different patterns displayed by the two categories in the results of the GOE-rating questionnaire. While verbal converses behave in a similar way to other adjectival and verbal opposites, nominal converses do not seem to fit into the same pattern, that of the distribution along a continuum of antonym canonicity.

**Table 3.7** *Nominal Converses (by overall GOE-rating)*

| Word 1 | Word 2 | GOE (1/2) | GOE (2/1) | GOE Overall | T-Score | EAT (1/2) | EAT (2/1) |
|---|---|---|---|---|---|---|---|
| employer | employee | 2.95 | 2.75 | 2.85 | 10.39 | n/a | n/a |
| pupil | teacher | 2.90 | 3.40 | 3.15 | 11.40 | 0.23 | 0.11 |
| husband | wife | 3.10 | 3.25 | 3.175 | 28.01 | 0.85 | 0.46 |
| parent | child | 3.10 | 4.08 | 3.59 | 15.66 | 0.25 | 0.02 |
| victim | murderer | 3.15 | 3.30 | 3.225 | 3.11 | n/a | 0.01 |
| master | apprentice | 3.20 | 3.25 | 3.225 | 2.17 | 0 | n/a |
| doctor | patient | 3.25 | 3.50 | 3.375 | 9.76 | 0.04 | 0.22 |
| landlord | tenant | 3.68 | 3.25 | 3.465 | 18.37 | 0.10 | n/a |
| student | teacher | 3.73 | 3.38 | 3.555 | 9.46 | 0.02 | 0.02 |

As Table 3.7 shows, all nominal converses have a GOE rank between 2.85 (*employer:employee*) and 3.59 (*parent:child*). *Husband:wife* has the highest EAT score of all pairs in this study and also scores as one of the highest pairs on the t-score. Nevertheless, the GOE-rating score, at 3.175, is not very high at all. *Apprentice:master*, both lexemes which are much less frequent than those in the previous pair, scores similarly on the GOE-rating questionnaire but only achieves a t-score of 2.17, barely above chance co-occurrence, and no scores on the EAT task.

The two graphs in Figure 3.10 show clearly that the rather erratic t-scores in this group are not reflected in the GOE-rating scores. Even in pairs with very high t-

[16] This applies to the adjectives in this study and may be a tendency observed throughout the lexicon but complexity of category structure is not exclusively tied to word class.

scores, like *husband:wife* and *landlord:tenant*, the GOE-rating does not reflect the strong associative bond of the two lexemes. The other striking observation illustrated by the GOE-rating graph above is the difference between this cluster of opposites and all others in this study: in all other graphs, a gradual increase in GOE-rating was observed whereas in Figure 3.10 above the graph with the diamond markers is almost level throughout the group.

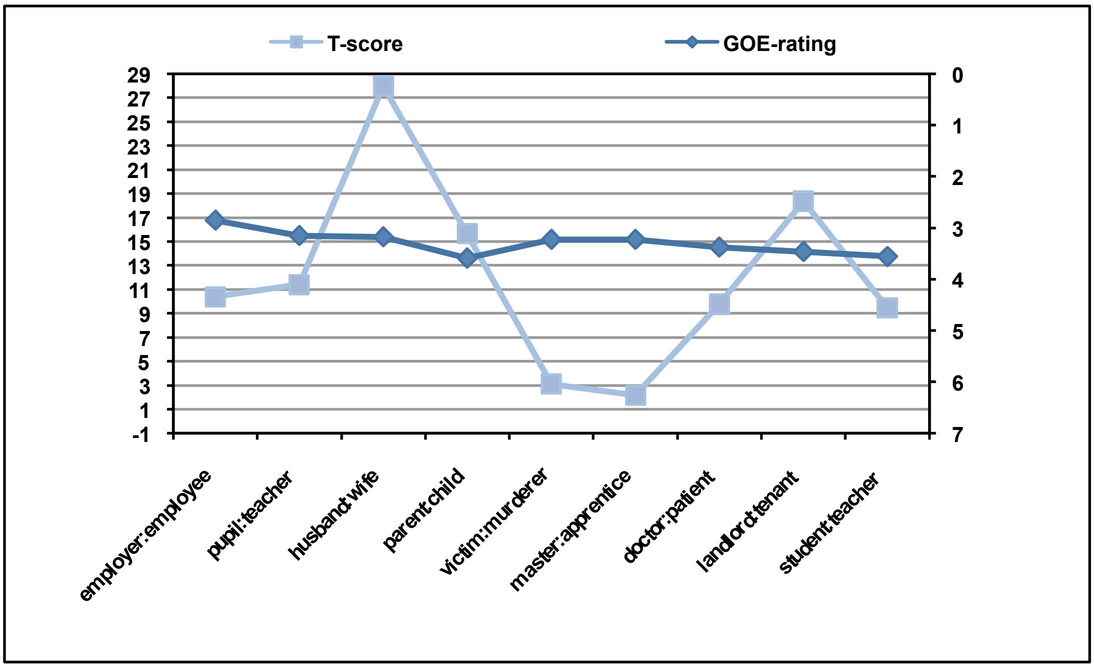



**Figure 3.10** *GOE-rating and t-scores for nominal converses*

The differences between the t-scores and the EAT scores of the nine nominal converse pairs are similar to the differences between those of the verbal converses. However, the GOE-rating scores are much more similar than in any of the groups discussed previously. One reason, as mentioned above, could be the difference in word class between the examples. Almost all pairs discussed in 3.4.1 were adjectival (with the exception of *giant:dwarf* which showed extremely high GOE-ratings for a noun pair) and therefore this phenomenon could not be observed in previous discussions. If goodness of antonymy judgements are indeed tied to word class there will have been very few studies in which this could have been investigated, as most research on antonyms has been conducted solely on adjectival pairs.

The theory which has been put forward in the previous section will be investigated in detail in 3.4.3 below where this difference between nouns and adjectives is also apparent. However, whether this difference is purely based on the cognitive properties of certain word classes remains to be examined.

Nominal converses do not seem to display the same pattern seen in the results for verbal and adjectival opposites. However, whether this is the case because the pairs under investigation are converses or because they are nouns is a question which cannot yet be answered and further data and analysis is needed to come to a satisfactory conclusion (cf. 3.5 and 4).

To conclude this section on converses, and to answer the question of whether converses are opposites, the factors which were considered to be of importance in this section will be summarised. One additional factor which could be contributing to judgements of opposition is word class. This has not been considered a criterion for antonymic strength so far and the following chapters will investigate the amount of influence this factor has on antonym judgements and the reasons for this influence. However, there are also striking similarities between converses and other types of opposition, especially in the category of verbal converses. Similar criteria for good opposition are applied both to converses and to other opposites and the distribution along a continuum holds for both classes. Most of the evidence displayed in this section indicates that, overall, converseness should be considered part of the category of opposition.

The question of converses will be re-examined in Chapter 5, which will look at differences in English and German converse judgements. The contrastive angle provides a wealth of evidence, especially in terms of the importance morphological relatedness plays in these judgements, as in German there is a prefix which indicates converseness (among other things): *ver-* (e.g. *kaufen:verkaufen*).

### 3.4.3 Complementarity: a gender issue (*MALE:FEMALE*)

The last cluster of pairs from the GOE-rating questionnaire which will be discussed in detail is that of opposite pairs based on the *male:female* distinction. *Male:female* is usually cited as one of the best examples of complementary opposites (e.g. Lyons 1977, Cruse 1986, Lehrer 2002) and it is therefore assumed to be part of the category of canonical antonymy. The GOE-rating questionnaire results are thus quite surprising as *male:female* does not fall within the group of excellent antonyms (scores of 1 to 1.8) but receives an overall rating of 2.06. The individual rating for the sequence *male:female* was, at 1.88, marginally higher, whereas the dispreferred sequence, *female:male*, received 2.18. Other pairs, such as *man:woman* and *mother:father,* were expected to score higher than they did, especially if t-score and results of the EAT task of these pairs are taken as a predictor of antonymic strength.

As Table 3.8 shows, the pairs in this group ranked among the highest in terms of frequency of co-occurrence as well as EAT scores. All but two pairs display reciprocal associative strength, with many of the scores reaching 0.50 or higher, which means that over 50% of the informants in the association task responded to a given word 1 with word 2 of the pair. However, as has already been mentioned, associative strength and frequency of co-occurrence alone are not satisfactory as indicators of the strength of the antonymic relation between two lexemes.

The discrepancy between the GOE-rating scores and the measures of associative strength which can be seen in Table 3.8 below provided a useful starting point for the investigation of other factors which could be influential in antonym judgements. After a detailed analysis of the results, a hypothesis will be put forward; this will then be examined more closely in Chapters 4 and 5 with the help of additional data (lexical decision tasks and German data) to ensure that it holds in cases other than this very specific subset of the category of opposites.

**Table 3.8** *Opposite pairs clustered around* male:female

| Word 1 | Word 2 | GOE-rating | T-Score | EAT (1/2) | EAT (2/1) |
|---|---|---|---|---|---|
| male | female | 2.06 | 32.54 | 0.64 | 0.49 |
| masculine | feminine | 1.81 | 8.94 | 0.47 | 0.27 |
| man | woman | 2.11 | 28.67 | 0.67 | 0.60 |
| husband | wife | 3.175 | 28.01 | 0.85 | 0.46 |
| mother | father | 4.55 | 29.99 | 0.52 | 0.54 |
| aunt | uncle | 4.525 | 10.15 | 0.62 | 0.58 |
| mum | dad | 3.5 | 26.47 | 0.70 | 0.56 |
| nephew | niece | 3.96 | 3.60 | n/a | n/a |
| brother | sister | 3.415 | 17.43 | 0.69 | 0.63 |
| king | queen | 4.775 | 15.23 | 0.45 | 0.40 |
| cow | bull | 3.755 | 2.18 | 0.11 | 0.38 |
| actor | actress | 4.4 | 2.96 | n/a | 0.24 |

The previous section on converses has shown word class to be a possible factor in antonym judgements. This is a hypothesis which will be investigated further here as, once again, the data displays a difference in GOE-scores between adjectival pairs and nominal pairs. However, in this case, the difference is not as substantial and one nominal pair, *man:woman*, receives an almost identical score to *male:female*. Although *male:female* is commonly considered the base pair, *masculine:feminine* is the pair with the best GOE-rating result despite a much lower t-score and lower EAT results.

Figure 3.11 below shows that despite several very high t-scores in this group of opposite pairs, the GOE-ratings are extremely variable and, once again, there is no direct correlation between the two measures. The data analysed in this section does not lend support to the claim that a word pair which has antonymic ‘potential’ becomes entrenched as an antonym pair through frequent co-occurrence. It seems a more likely hypothesis that the frequent co-occurrence of two members of an antonym pair is, as previously mentioned (cf. 2.5.3), a consequence or by-product of their shared semantic range and the larger overlap in semantic features which makes these lexemes more likely to appear in the same contexts.

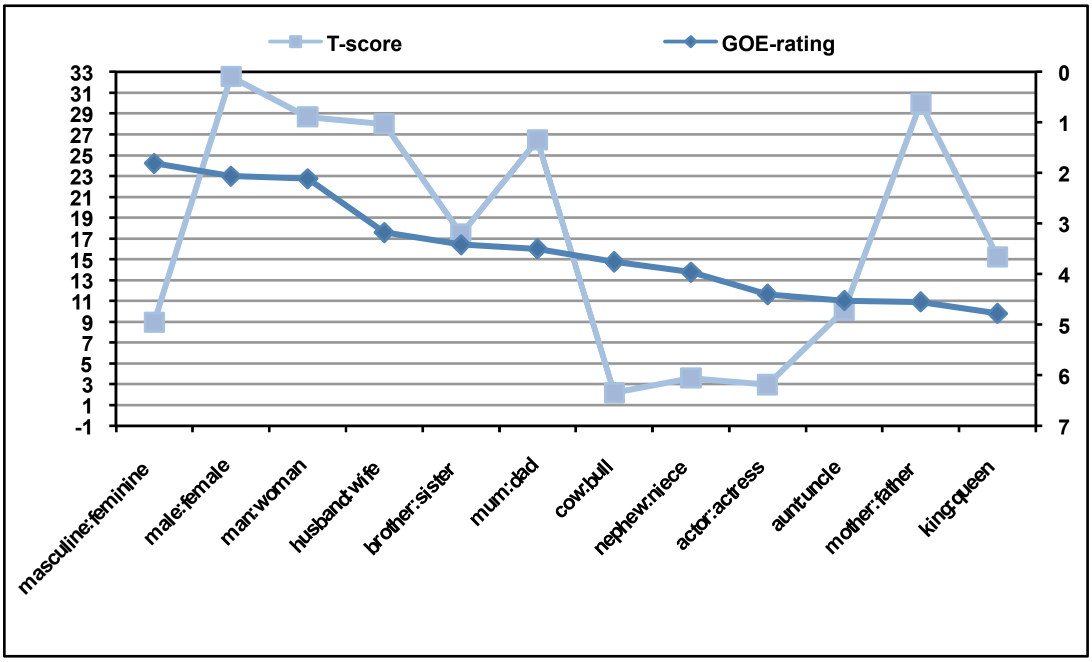



**Figure 3.11** *GOE-rating and t-scores for pairs around* male:female

The first group under discussion is that of the six family term pairs: *husband:wife*, *mother:father*, *aunt:uncle*, *mum:dad*, *nephew:niece* and *brother:sister*. Of these six, only the first three were included in the original questionnaire but the fact that this category seemed to provide interesting results led to the inclusion of the remaining three pairs (and two others which were not considered opposites: *father:daughter* and *mother:daughter*). The three highest scores in this group are those of *husband:wife* (3.175), *brother:sister* (3.415) and *mum:dad* (3.5). None of these pairs are considered better than ‘medium’ from the point of view of antonymic strength. There is, however, one possible explanation for the comparatively high scores of *husband:wife* and *brother:sister*. Both these pairs stand for reciprocal logical relationships between the concepts the individual lexemes encode. Logically, if A is B’s husband, B is A’s wife. By the same token, *mother* should be paired with *child*, *son* or *daughter* (as a converse) and not with *father* as there is no logical relationship which holds between these two concepts; this also applies to *aunt:uncle* and *nephew:niece*. However, *mother:daughter* and *father:daughter* (the logically ‘correct’ combinations) score even lower than *mother:father* in the GOE-rating (5.08 and 5.415 respectively) despite relatively high t-scores (15.33 and 9.96). Before a hypothesis can be put forward which might provide an explanation for the above scores, the remaining pairs will be discussed briefly.

None of the last three pairs, *king:queen*, *cow:bull* and *actor:actress*, receives a high score and while the t-score for *king:queen* is very high, the other two pairs barely reach a statistically significant t-score. However, it is between these two pairs that the greatest difference can be seen. *Cow:bull*, with a GOE-rating score of 3.75, is the highest scoring of this group whereas the other two pairs are significantly lower with scores far below four. While all of these are nouns, *cow* and *bull* are not lexemes which call up particularly rich cognitive models as, especially for people living in Western urban society, they are not particularly salient concepts or ones which have an impact on day-to-day life. While biological gender is the only difference between a cow and a bull, there are other differences which result from the animals’ respective gender which are more salient than gender alone, i.e. ‘a cow gives milk’, ‘a bull has horns’ etc. These additional factors might contribute to a lower GOE-rating score as they distract from the salience of the *male:female* distinction and antonymy is a relation based on the principle of minimal difference (cf. Murphy 2003).

### A HYPOTHESIS: CONCEPTS AND IDEALISED COGNITIVE MODELS (ICMS)

One hypothesis which provides a satisfactory explanation of the much lower scores of the pairs above is strongly tied into the argument concerning word class. In cognitive linguistic terms, nouns are lexemes which encode cognitive categories associated with rich cognitive models which underpin these categories. (Idealised) cognitive models ((I)CMs) structure knowledge (linguistic and non-linguistic) which is associated with a certain cognitive category (cf. among others Lakoff 1987 and Ungerer & Schmid 2006). They furthermore generate prototypes as by-products of this structuring process. These models do not usually reflect facts in the real world but the interpretation and experience of an individual and are strongly influenced by culture. Thus they also include encyclopaedic knowledge which will then, undoubtedly, influence not simply decisions of category membership but also other decisions related to the associations formed by the concept and by extension the lexeme which encodes it (i.e. antonymic strength among others).[17]

---

[17] For a detailed discussion of ICMs, see Lakoff 1987 and for an overview Ungerer & Schmid 2006.

Adjectives encode conceptual categories which do not usually evoke rich ICMs, and the ICMs of verbs are usually less well developed than those evoked by nouns[18]. This is reminiscent of Cruse's criterion of purity of opposition, as a rich ICM certainly involves several bundles of attributes which would make the concept and the lexeme 'impure'. The lack of a rich ICM for adjectives, however, makes them perfect candidates for a relation which is, on the one hand, determined by minimal difference and on the other requires this minimal difference to be the most salient aspect of the meaning of both lexemes.

To illustrate the above hypothesis, one pair, *mother:father*, will be investigated in more detail. This pair scored very low on the GOE-rating (4.55), especially considering it is one of the pairs with the highest co-occurrence in the entire study (29.99). The fact that there is no direct logical relationship between the two concepts has already been discussed and will, for the moment, be disregarded.

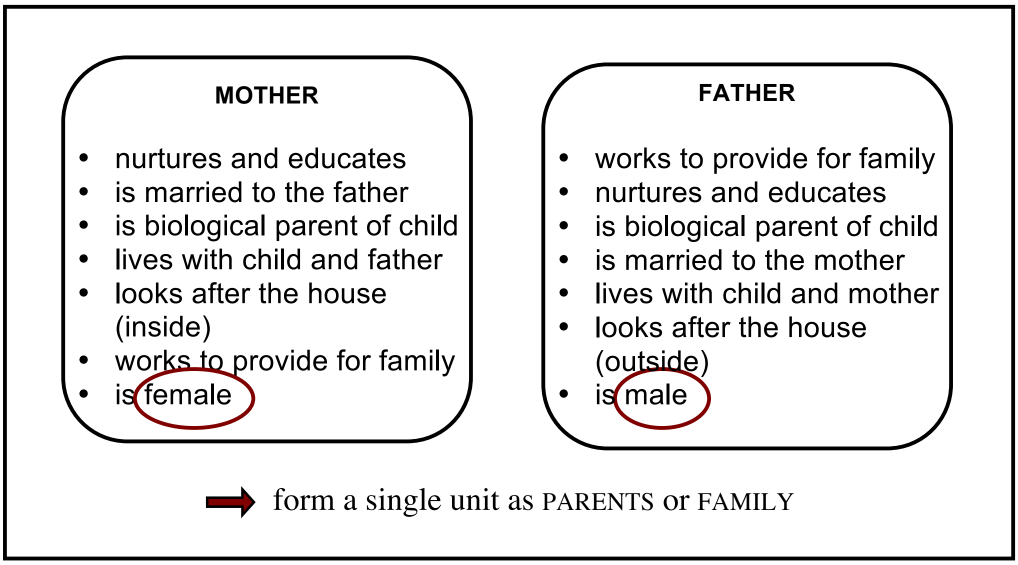



**Figure 3.12** *Possible overlapping* MOTHER *and* FATHER *ICMs*

Figure 3.12 shows a highly idealised (and stereotypical) cognitive model of the concepts of both MOTHER[19] and FATHER. It can be seen that the attributes are largely the same for both MOTHER and FATHER, albeit in a slightly different ordering. There are, of course, more attributes and the list could be extended almost infinitely. However, some of the most salient ones are listed below and these demonstrate clearly that the two concepts have many attributes in common. This is, at least not

[18] This is not the case for all adjectives (nouns and verbs) but happens to be true for those concepts under closer investigation here. The complexity hypothesis would also predict that more complex adjectival and verbal concepts, for example, will be less easily matched with their opposite than nominal ones which rely on a less complex category structure.

[19] As this model is, at this stage, only intended for illustration, I have not taken into account Lakoff's (1987) complex ICM of MOTHER.

initially, a concern for antonymy as it is based on minimal difference. However, it seems essential that the salience of this difference be greater than that of the other attributes and this is clearly not the case in the model below (cf. Figure 3.12).

It is debatable whether, in an attribute listing task (as in Rosch & Mervis 1975) of the concepts in question, 'male' and 'female' would even be listed as attributes of MOTHER and FATHER.[20] It is certainly not expected that they would occur at the top of the list and as all attributes in cognitive categories or models are weighted, the weighting of 'male' and 'female' as attributes of MOTHER and FATHER would receive a very low weighting indeed.

This theory might also go some way towards explaining the better scores of *man:woman* and *mum:dad*. *Man:woman* does not have particularly rich ICMs associated with either lexeme and in an attribute listing task there would most likely be significantly fewer attributes weighted more strongly than 'male' or 'female' than in the case of *mother:father*. *Mum:dad*, a pair which was originally included to investigate the possibility of emotive aspects playing a role in antonym judgements, evokes ICMs which would presumably look like an impoverished version of the FATHER and MOTHER ICMs, as *mum* and *dad* are not simply the colloquial synonyms of *mother* and *father.*[21] The use of these more 'familiar' terms foregrounds certain aspects of the PARENT ICMs while others are relegated to the background or even lost. This might result in increased importance of the *male:female* distinction. However, this last hypothesis is, at this point, merely speculation.

One of the nominal pairs which scored extremely high on the GOE-rating, *giant:dwarf*, displays evidence for the opposite to the phenomenon above: the fact that certain pairs differ along more than one salient axis might well support their antonymic strength. However, another factor specific to this pair is that the attribute which is the most salient and would almost certainly be the most strongly weighted in the case of *giant:dwarf* is that of 'size'. Therefore it is unsurprising that this pair, despite being nominal, has scored relatively highly. This hypothesis will be revisited in Chapter 6 and will also be developed further with the help of additional data analysis in the following two chapters.

To conclude this section, returning the surprisingly low score received by *male:female* in the GOE-rating questionnaire, the theory above can also apply to adjectival pairs. Whereas most adjectival pairs do not evoke strong ICMs, there are many pairs which call forth a certain amount of associative 'baggage'. Considered

[20] The results of such a task are discussed in 3.4.4.

[21] Register and jargon can both have a significant effect and some pairs will only be considered opposites in a particular linguistic context. Paradis et al. 2007 found that one of the most successful pairs in their corpus analysis was the medical *open:laproscopic* referring to types of surgery (cf. 6.1.3).

only from a biological perspective, *male:female* should be excellent opposites. However, English conflates sex and gender in the lexemes *male* and *female*. As there are no separate words to refer solely to biological gender, this is mixed together with 'social' gender and therefore carries many connotations and additional encyclopaedic knowledge which is not usually the case with other adjective pairs. Despite the fact that this is not a completely satisfactory explanation, this and the discussion above goes some way towards explaining the peculiar discrepancies found in the data for this cluster of opposites.

### 3.4.4 Attribute listing task

In order to confirm the hypotheses presented above, a short attribute-listing task (cf. Rosch & Mervis 1975) containing some of the concepts discussed above was carried out. The first hypothesis was put forward as a result of the unexpectedly low score of some of the pairs based on *male:female*. It states that antonym pairs which consist of concepts where the antonymic attributes do not play a central role in the category-internal structure will score lower than those where the opposed features have greater salience. In the structure of the concepts of MOTHER and FATHER, for example, the attributes *male* and *female* are not as central as other features. The characteristics *male* or *female* are superseded by other attributes which are considered more salient and are thus weighted higher and this weakens the opposition since, while antonymy is a relation based on minimal difference, this difference should be in the most salient attribute of the members of an antonym pair.

The second hypothesis concerns the density and complexity of the concepts involved in an antonymic relation and suggests that the more complex the concepts are, the less likely they are to form a very strong antonymic relation. This may be an explanation for the reason why some word classes, for example adjectives and prepositions, are generally considered better antonyms than others. The data presented below will be used to build the structure of a pair of cognitive models (MOTHER and FATHER) and will, based on the discussion of the antonym pairs *mother:father* and *man:woman* in 3.4.3, take a closer look at the validity of the two hypotheses.

#### *3.4.4.1 Methodology*

The attribute listing task was designed in line with Rosch & Mervis (1975) and was divided into Task A and Task B. Task A contained 15 nouns which were one member of one of the opposite or control pairs used in the GOE-rating questionnaire and Task B was composed of the other member of each of the pairs. The word pairs used in the attribute-listing task are listed in table 3.9 in the order in which they were presented to the participants.

**Table 3.9** *Concepts used in the attribute listing task*

| TASK A | TASK B | TASK A | TASK B | TASK A | TASK B |
|---|---|---|---|---|---|
| carrot | pea | dog | cat | giant | dwarf |
| cow | bull | employer | employee | rain | wind |
| apprentice | master | glove | hand | woman | man |
| coffee | tea | father | mother | valley | hill |
| king | queen | mug | cup | student | teacher |

Each participant was thus only presented with one member of a pair to avoid a synchronisation of the responses. Participants were asked to list as many attributes and characteristics of the given concepts and objects as possible in one minute. As the task was completed online without supervision of any kind, the participants were asked to list the attributes in the order they had thought of them and not to spend more than sixty seconds on each item or to go back and change their responses. The survey tool used was, as explained in 3.2.1, an online questionnaire-design tool called SurveyMonkey. An example of the listing task can be seen in Appendix 7.

Each of the questionnaires was filled in by twelve informants between the ages of 19 and 34 in their own time. All participants were native speakers of a variety of English, but neither variety nor age was considered as a factor in the responses. The results were collected for each item and a complete list can be found in Appendix 8. Both tasks were also translated into German and fifteen participants filled in each of the two questionnaires containing the German equivalents of the English items. This data will be discussed contrastively in 5.2.4 and overall conclusions for canonicity and antonymic strength will be drawn in Chapter 6.1.3.

*3.4.4.2 Discussion and analysis of results*

Each item elicited between one and ten attributes and the average number of attributes listed overall was between five and six. In this section only the results for MOTHER, FATHER, WOMAN and MAN are discussed in detail and their attribute structure is compared to that of GIANT and DWARF. Further discussion of other items used in the task will follow in Chapter 6 where the above mentioned hypotheses are presented in their final, slightly modified form alongside the theoretical conclusions drawn from this short study, which also take the results of the behavioural data and the contrastive analysis of the English and German data into account.

Figure 3.13 shows the results for MOTHER and FATHER: only attributes which were mentioned more than twice in the listing task are included here and all characteristics are ranked by number of occurrences as well as position of occurrence (whether they were listed as the first or last attribute). Attributes which were listed first were weighted at eight, those listed second at seven, down to any attributes which were listed eighth and lower which were weighted at one. The numbers in brackets refer to the number of participants who listed the attribute in question and the overall weighting of this attribute. Thus the attributes at the top were listed first by the largest number of participants.

| **MOTHER** | |
|---|---|
| has children/gave birth | (10; 71) |
| is caring | (7; 46) |
| is loving/loves | (6; 41) |
| is warm/gentle | (7; 49) |
| nurtures/nourishes | (5; 33) |
| is female/woman | (4; 27) |
| comforts | (3; 23) |
| works/hard-working | (3; 17) |
| cooks | (3; 12) |
| listens | (2; 11) |
| teaches | (2; 4) |

| **FATHER** | |
|---|---|
| has children | (4; 31) |
| is a man/is male | (4; 30) |
| provides | (2; 15) |
| is old | (2; 14) |

**Figure 3.13** *Attribute listing results for MOTHER and FATHER*

Surprisingly, as can be seen in the figure above, the results for FATHER are markedly different from the results for MOTHER. Overall there were far fewer responses to FATHER, many of which were very subjective (e.g. ‘mine is great’), and very few attributes were listed by two or more participants.[22] However, it is clear from the listing that ‘male’ and ‘female’ are not very prominent as attributes of either concept. Both are mentioned by the same number of informants (4 in each case) with roughly the same weighting. The category MOTHER has a considerable number of attributes which are considered more noteworthy than ‘female’ whereas FATHER only has four attributes in total which were mentioned twice or more. This discrepancy between the results for MOTHER and FATHER highlights the complex nature of the category structure. Compared to the concepts used in many other attribute listing tasks, some of the categories in the present study are more difficult to define by attribute listing

[22] For a full list of responses, see Appendix 8.

since their key features are largely abstract and thus more difficult to define. Many of the features listed for MOTHER refer to the stereotypical and idealised role of a mother in the family unit whereas the role of the father as provider is less clear from the responses.

These results seem to support the first hypothesis put forward above, namely that the salience of the antonymic features (in this case ‘male’ and ‘female’) is crucial to the construction of a pair of lexemes as an antonym pair. Since the features in question do not rank particularly highly in the attribute-listing results and are thus considered less central to the overall structure of the category, *mother:father* would not be predicted to be a particularly good opposite pair. Furthermore, looking at the second hypothesis, the difference in complexity of the categories is problematic for antonym construal for two reasons: firstly, the lexemes in question are clearly not as minimally different as better antonym pairs would be and secondly, because the complexity of the category structure seems to reduce the weight of every individual feature thus further diluting the salience of the potentially antonymic features.

| **WOMAN** | |
|---|---|
| has children | (8; 42) |
| mother(hood) | (3; 22) |
| beauty/beautiful, | (3; 19) |
| female | (2; 16) |
| feminine/femininity | (2; 15) |
| breasts | (2; 15) |
| nurturer/nurturing | (2; 12) |
| skin | (2; 8) |
| smells good/nice | (2; 6) |

| **MAN** | |
|---|---|
| male | (7; 55) |
| strong/firm | (6; 41) |
| has a penis | (5; 31) |
| not woman/opposite of | (4; 29) |
| woman | (3; 20) |
| human | (4; 16) |
| testosterone | (2; 13) |
| taller than women | (2; 12) |
| mankind/humanity | (2; 11) |
| has two legs | (2; 10) |
| facial hair | (2; 9) |
| has two arms | (2; 8) |

**Figure 3.14** *Attribute listing results for WOMAN and MAN*

Figure 3.14 shows the listing results for *man:woman* and it can be seen even from a superficial examination of the table that at least the number of features is more balanced than in the previous example. While the attribute ‘male’ is at the top of the list for the category MAN, only two participants list ‘female’ as an attribute of the category WOMAN. In this case, the categories are reasonably well matched for complexity but nevertheless both are fairly complex. The antonymic features are slightly more central to the category structure and thus, simply on the interpretation of the data along the two hypotheses presented above, *man:woman* would be expected to form a better antonym pair than *mother:father* where more other, non-antonymic features are in focus in both members of the pair.

The last example under discussion is the opposite pair *giant:dwarf*. The results in Figure 3.15 show a considerably different result from those in figures 3.13 and 3.14. Neither of the two concepts has a large number of features and the features are very well matched across the two categories. If ‘tall’ and ‘big’, as well as ‘short’ and ‘small’, are combined to form one feature each, neither category has more than five features which were listed by two or more participants. Furthermore, the two categories have not just one set of antonymic features (*big:small*) but two (*short:tall*) in the first three attributes listed. The first feature was in both cases a size adjective (*big/large* or *small*) in all listings but one.

This example, which is rated as a much better antonym pair than either *man:woman* or *mother:father* in the GOE-rating, shows the importance of the salient position of the antonymic features in determining how good an opposite pair two concepts will be. The overall category structure of the concepts GIANT and DWARF is also significantly simpler than that of the other concepts under discussion in this section. Both these factors contribute to the pair’s higher score in judgement tasks and, most likely, also in the behavioural study discussed in the following chapter.

| **GIANT** | |
|---|---|
| huge/big/large | (12; 83) |
| myth/fairytale | (7; 44) |
| tall | (4; 26) |
| has two legs | (2; 13) |
| man/male | (2; 12) |

| **DWARF** | |
|---|---|
| short (not tall -1) | (11; 87) |
| small (stunted - 1) | (9; 70) |
| Snow White | (7; 43) |
| wears (funny) hat | (4; 24) |
| beard | (3; 19) |
| exaggerated features | (2; 14) |

**Figure 3.15** *Attribute listing results for GIANT and DWARF*

Overall, it can be said that both hypotheses put forward at the beginning of this section are supported by the data discussed here and go a long way towards explaining the peculiar patterning observed in the analysis of the GOE-rating data for nominal opposite pairs. Especially if it is the case that non-canonical pairs are processed ‘online’ rather than stored together and recalled as a ‘chunk’, it stands to reason that pairs with a simpler overall category structure and those with the antonymic features in the most salient position will be more readily recognised as opposites. This section has, as mentioned above, only dealt with a small number of examples to illustrate the validity of the hypotheses put forward on the basis of the GOE-rating data discussed in 3.4.3. The attribute-listing results will be discussed further in 5.2.3 where the English data will be compared to that obtained from German native speakers to investigate cross-linguistic differences. In Chapter 6.1.3, both hypotheses will be re-evaluated using not only the GOE-rating and attribute-listing data but also the behavioural data discussed in Chapter 4.3 to draw theoretical conclusions and to determine how reliable an indicator complexity of category-internal structure and salience of antonymic features are of antonymic strength.

## 3.5 Conclusions

The aim of the analysis above was twofold: firstly, to investigate whether associative strength, indicated by data from a free word association task (EAT) and frequency of co-occurrence data from the BNC, is a reliable precursor for antonymic strength (GOE-rating results). If, as many researchers claim (cf. among others Miller et al. 1989, Charles & Miller 1990, Justeson & Katz 1991), canonical antonymy is a purely associative relation which is entrenched by high frequency of co-occurrence of the two members of an antonym pair (in certain syntactic frames), antonymic strength should be predictable by consulting frequency of co-occurrence data and other measures of associative strength for each pair. However, as the data above has shown very clearly, this is not the case. There are, of course, instances where the associative strength and antonymic strength of a particular pair match, but whether the reason for this is that frequency of co-occurrence is the main determining factor remains unclear. This match between antonymic and associative strength occurs mainly with strongly canonical antonym pairs (e.g. *hot:cold, large:small, good:bad*). As soon as the pairs under investigation are taken from outside this central category of canonical antonymy, however, it becomes clear that associative strength can no longer be depended on as a reliable indicator of antonymic strength. In these antonym judgements other factors seem to play a more dominant role. The investigation of these factors was the second aim of this chapter. Several criteria for good antonymy, which had been introduced in Chapter 2.5, were revisited here and their explanatory power for the phenomena observed in the data was considered. Furthermore, a number of new criteria were introduced and evaluated.

Before concluding this chapter, I would like to briefly re-examine all the relevant factors in turn and evaluate the role they played in explaining the data gained from the goodness-of-exemplar rating questionnaire. There are eight factors which were found to influence antonymic strength; one, **frequency of co-**

**occurrence** (or breadth of co-occurrence) has already been discussed above. Suffice it to say here that, at present, it is not clear whether high frequency of co-occurrence is indeed a predictor or whether it is a consequence of particularly high antonymic strength. It is, however, evident that there is a strong connection between associative and antonymic strength, especially where canonical opposites are concerned.

Another influential factor, but one which has not been discussed extensively in this chapter, is **morphological relatedness**. Whether the two members of an antonym pair share morphological material is a very important factor in antonym judgements. The lowest scoring morphologically related pair in the data is the nominal pair *tutor:tutee* (3.39) and the lowest adjectival pairs are *married:unmarried* (1.965) and *interested:disinterested* (2.075). This alone shows that morphologically related pairs receive very high ratings, not simply because of the shared morphological material but equally because of the implicit opposition expressed through prefixation. This is one factor which will be re-examined in greater detail in the section which investigates differences and similarities in antonym judgements in English and German (Chapter 5).

One element which proved to be particularly important in the judgement of scalar antonyms is that of **symmetry** (Cruse 1986 – 'equidistance from the mid-point of the scale'). The analysis of *hot:cold, good:bad* and the antonyms on the size scale, as well as the peripheral pairs of the same fields, has shown that pairs which are located equidistant from the mid-point of the scale which structures the antonymic domain consistently score higher than pairs which are asymmetrically placed. This can be traced back to the criterion of minimal difference as, if the distance to the mid-point is identical, this reduces the differences in range and distribution significantly.

However, considering the distribution of the members of antonym pairs along a scale, symmetry is not the only factor which is seen to play a role. The **overall distance** between the concepts the lexemes encode is also of importance. It seems that lexemes which are at a greater distance along the scale, e.g. *excellent:bad* as opposed to *bad:satisfactory* or *good:mediocre*, are rated as better antonyms than those which cover parts of the scale which are closer together.

Cruse's (1986) second criterion, **purity of opposition**, also seems to play a role. It is unclear, however, how closely this is linked to the newly introduced factors of **word class** and **associated material** ('baggage') which are part of the meaning of the lexemes of a pair and thus influence antonymic strength. It could be that purity of opposition (in Cruse's sense) is a sub-category of this wider approach and will therefore no longer appear as an individual factor. The idea that different lexical categories are associated with different conceptual structures is by no means new (cf. e.g. Langacker 1982) and in cognitive semantics it is generally assumed that the structures of nominal categories on the one hand and verbal and adjectival categories on the other are quite different. This is evident in the data and might, if it holds up under closer investigation, provide an explanation for the fact that nominal opposite pairs generally score much lower on antonymic strength ratings than their verbal and adjectival counterparts. Different category structures with different amounts of semantic and connotational material contained in the word meaning potentially weakens the otherwise salient antonymic axis opposites depend on.

The last criterion, which has previously been put forward and examined by both Cruse (1986) and Muehleisen (1997), is that of **semantic range**. The data clearly shows that semantic range plays an important part in antonym judgements and that this criterion is another factor related to frequency of co-occurrence as a larger amount of shared semantic range will provide the lexemes in the pair with more opportunities to co-occur and therefore strengthen their associative relationship. However, shared semantic range also presupposes a large amount of overlap in meaning as well as factors such as register which, in turn, corresponds again to the demand of antonyms for minimal difference.

Having examined all the above criteria and tested their validity on the data, the next step is to substantiate the evidence that these factors indeed contribute to antonym judgements and attempt to separate them into primary and secondary factors as well as trying to ascertain which of the criteria are reliable predictors and which are merely consequences of the interplay between other factors. The following chapter will attempt to clarify these points by analysing data from two lexical decision tasks as well as using statistical methods to determine the most prominent predictors.

Since it cannot be taken for granted that the results of the judgement task discussed in this chapter are representative of the structure of lexical storage and the degrees of antonymicity in the mental lexicon, the analysis of behavioural data will aid in determining what can be concluded about the actual representation of antonymic relations. This is a question which has been much discussed by previous research, and the judgements emerging from the experiments reported on in Chapter 4, which were made with less conscious deliberation due to the online design of the tasks involved, will help determine whether the judgements discussed above constitute substantial evidence for the representation of opposites in the mental lexicon.

# 4. Lexical decision tasks

The opposite of *doctor*? Well,
That's not so very hard to tell.
A *doctor*'s nice, and when you're ill
He makes you better with a pill.
Then what's his opposite? Don't be thick!
It's *anyone who makes you sick*.
(Wilbur 2004: 519)

This chapter presents two psycholinguistic experiments and their results which are discussed individually. The experiments are designed in a similar fashion to standard lexical decision tasks and require the participants to distinguish between antonymic and non-antonymic word pairs presented to them. First, Experiment 1 and its results will be presented before bringing this data together with the measures of associative and antonymic strength discussed in Chapter 3 in order to draw conclusions about lexical opposition and the extent to which these relations are entrenched in the speakers' mental lexicon or whether they are, at least to some degree, constructed 'online' through, for example, a process of feature matching (cf. Hutchison 2003). Certain insights into how the mental lexicon might store polysemic lexemes and whether lexemes have a preferred antonymic partner which result from the second of the two experiments will be discussed separately at the end of the chapter.

## 4.1 Theoretical considerations

The experimental methods used for data collection as well as the results discussed in this chapter are based on several theoretical concepts and mechanisms which will be introduced briefly below. A vast body of research and a large number of theories exist on the subject of the organisation of lexical storage and the means by which this knowledge is accessed, not just in everyday conversation but also in experimental tasks. For the present study, it is not necessary to give a complete overview of the field but it is essential to illustrate (through the use of a select number of examples) the theoretical foundations on which the experimental research and its analysis are based.

### 4.1.1 Spreading activation and semantic priming

The first concept to be discussed is that of **spreading activation**, since it forms the basis of the experimental method used in the present research. The mechanism of spreading activation is considered the dominant in theories providing an explanation for the phenomenon of semantic and associative priming (cf. Harley 2008: 190). It is based on decades of experimental research including early (and later) word association and elicitation experiments (e.g. Deese 1964 & 1965, Paradis et al. 2007 & 2009) and the psycholinguistic methods mentioned above. The principle underlying spreading activation is relatively simple: it is assumed that from any entry point into the lexicon – a visual, auditory or other sensory stimulus - activation spreads along connections between nodes of a semantic network and activates nodes which are directly and indirectly linked to the entry node to greater and lesser degrees (cf. among others Collins & Loftus 1975, McNamara 1992). Activation spreads over time and decreases with increasing distance from the entry node (cf. McNamara & Altarriba 1988, McNamara 1992). A simplified illustration of activation spreading through a network is presented in Figure 4.1 below.

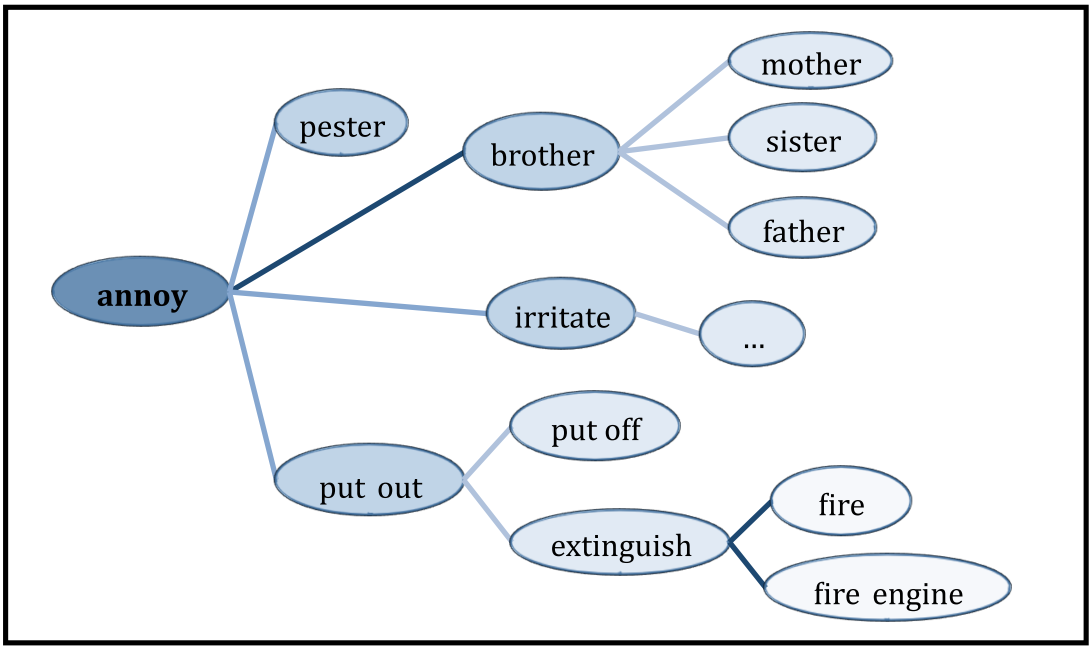



**Figure 4.1** *Illustration of the mechanism of spreading activation*

The entry point, *annoy*, is most strongly activated and with each additional level of distance, the activation of the nodes decreases. The links in the illustration below are of two kinds: semantic and associative. The former are based on a semantic relationship between the two nodes (synonymy, antonymy, super- or subordination – e.g. *pester, irritate, put out*) whereas the latter are purely associative (e.g. *brother*) and can be highly individual in nature. The association between two items can be based on meaning (in the case of semantic priming), or on structural, graphemic or phonological criteria, or on frequent co-occurrence and subsequent entrenchment of the combination of lexemes. Two items are considered associatively related if one is produced as a reaction to the other in a word association experiment (cf. McNamara 1994, Harley 2008). Spreading activation underlies a large number of models of lexical storage and access, among others all connectionist models. Some selected models will briefly be introduced in 4.1.2.

Spreading activation, as a mechanism which underlies semantic priming, is based on a very large body of research and priming is a widely used experimental method in conjunction with lexical decision tasks. The lexical decision task (LD task) is one of the most common experimental designs used in psycholinguistics. In a LD task, the participant is asked to decide whether a word presented to them (either visually or auditorily) is a word of a given language (in the case of the present study, English or German) and has to indicate their response by pressing a yes/no button. The data recorded during this experiment consists firstly of the participant's reaction time and secondly of any errors which were made during the experiment (e.g. 'yes' response to non-words). The time it takes participants to react to a single item is usually between 500 and 1000ms (cf. Harley 2008). A slower reaction time is associated with higher levels of processing effort and will thus lead to the assumption that the item in question is more difficult to access and/or process.

There are some potential problems with this design; one of them is a speed-accuracy trade-off (cf. among others Pachella 1974), as it has been noted that the faster participants are asked to react, the less accurate their responses are likely to be (and vice versa). However, the benefits of this type of task outweigh these kinds of drawbacks. Several factors, for example frequency, familiarity and number of letters and syllables, affect the reaction time in lexical decision tasks and need to be carefully controlled for. The non-words used as control items have to match the target words in the above properties as closely as possible to reduce the number of variables between the stimuli and allow for accurate statistical analysis.

Priming, a technique often used in LD tasks, originated in the late 19th century (cf. Cattell 1947). In priming, the participants are presented with a first stimulus (the ***prime***) which can be, for example, a word, sentence (presented visually or auditorily) or picture. Then after a certain interval which can be manipulated to fit the questions under investigation the participants are presented with a second stimulus, the **target,** which they react to by pressing a response button. The prime will influence the reaction time to the target and can either have a facilitatory or an inhibitory effect. Due to spreading activation, the reaction time to the target will be shorter if the prime is either semantically or associatively related to the target (e.g. *doctor-nurse*). Depending on which combination of modalities is used to present the primes and targets, LD tasks can be either intramodal (visual-visual, auditory-auditory) or cross-modal (auditory-visual or visual-auditory). The task used in the present study is described in more detail in 4.2.1 after a brief summary of models of lexical storage and access and an overview of investigations into the importance of antonymy in the structure of the mental lexicon.

### 4.1.2 Models of lexical access and the mental lexicon

The mental lexicon is most commonly thought of as a type of network consisting of nodes (concepts) and links of different types and varying degrees of strength. There is still considerable debate among experts about how we access this lexical knowledge, what precisely is stored in the lexicon and how this 'storage unit' is structured.[1]

Quillian (1962, 1967) developed a theory of semantic memory which was intended to assist in computer simulation of the process involved in semantic memory search and priming. This theory was extended by Collins & Loftus (1975) in order to have greater explanatory power for psycholinguistic experimental results. Quillian's theory 'viewed memory search as activation spreading from two or more concept nodes in a semantic network until an intersection was found' (Collins & Loftus 1975: 407). Despite the fact that both the original and extended version of the theories are geared towards explaining data resulting from experiments which use a priming paradigm, there are useful generalisations here which can be applied to the present research.

Quillian's notion of concepts and their representation as nodes in a network is still very much in evidence in more current research (cf. Aitchison 1994, Harley 2008). In his theory (cf. Quillian 1962, 1967), concepts correspond to particular senses of words or phrases (e.g. *machine*, *driving a car*, *the old car I own*). These concepts have a complex structure and build vast, open-ended interlinked networks.[2] According to Collins & Loftus (1975), each concept can be represented

---

[1] Since the tasks in the present study are ones which rely on lexical retrieval and activation rather than production, models of lexical production will not be included in the discussion.
[2] This can also be found in the theory of cognitive models in current cognitive linguistics (cf. Ungerer & Schmid 2006). The term **cognitive model** is used to cover "all the stored cognitive representations that belong to a certain field" (Ungerer & Schmid 1996: 47). Thus, a cognitive model can be seen as containing not only categories but also frames and scripts (cf. Fillmore 1985).

as a node in a network and its properties are represented as bi-directional relational links from the node to other concept nodes. These links are weighted (each link is assigned a *criteriality*) and the importance of a link to the concept depends on its weighting.[3] This way of conceptualising lexical storage and activation is common to many models of visual word recognition, especially connectionist models.

There are, broadly speaking, two types of models of how (visual) word recognition works: serial and parallel processing models. One of the most well-known examples of the former is Foster's (1976, 1979) **autonomous serial-search model** which proposes a two-stage process of lexical access and retrieval: the first step, when a lexical item is encountered, is a sequential search through modality-specific access files, which are subdivided by phonological or orthographic properties of lexical items and ranked by their frequency. The second step is the use of pointers from the located correct access file to the master file where all information about a particular word is stored. The first step has to be completed before the second stage can be initiated, hence the term 'autonomous.' In contrast to Forster's model, parallel processing models like Morton's **logogen model** (1969, 1970, 1979) and McClelland & Rumelhardt's connectionist **interactive activation model** (1981) only involve a single stage and rely on interactive processing as well as spreading activation which, in both models above, is cited as the explanation for priming. Whereas the logogen model operates with differing thresholds of activation for each logogen involved in the recognition process, and the activation threshold differs in terms of frequency, in the IAC model, Quillian's weighted connections make a reappearance as different connections are assigned different weights and the model further draws a distinction between **excitatory** and **inhibitory** connections. Connectionist models, common not just in psychological approaches to lexical storage but also in computational modelling of natural language, are mathematically complex and will not be discussed in detail.

To be able to assess what the experimental results in this study indicate, the question of the mechanisms behind semantic priming has to be addressed. Antonymy, as a lexical relation, falls into the category of semantic priming, which has been explained in the literature as being either purely due to associative strength (Harley 2008: 185ff.) or due to a matching of features of the two concepts during processing. Hutchison (2003), in a very thorough investigation of a large number of priming experiments and their results, concludes that both associative strength and feature overlap have an influence on semantic priming. What he calls

[3] For a more detailed discussion of Quillian's model, see Collins & Loftus 1975.

‘functional associates’ (Hutchison 2003: 785) – synonyms and antonyms – support a theory based on feature overlap whereas mediated priming paradigms seem to indicate that there is at least some level of automatic associative priming. In the analysis below, it can be assumed that automatic associative processing as a result of the entrenchment of close associates with great associative strength will play a role at the top end of the antonymic strength continuum. Opposite pairs which do not possess this added ‘bonus’ of associative strength, however, will be recognised by a process of feature matching, which is not quite in keeping with Collins & Loftus’s (1975) model of spreading activation since they stated that activation will automatically spread from any given node to all its associates, albeit at different speeds.

### 4.1.3 Antonymy and the mental lexicon

As far as the structure and mechanisms in the mental lexicon are concerned, Aitchison (1994: 95) succinctly sums up what is still considered the central question in research into lexical storage: ‘Do humans work things out, or look things up?’ And furthermore, if we do work things out, how do we do it and do ‘logical’ relationships play a role in the process of ‘working out’?

Word association experiments (e.g. Deese 1964 & 1965) have given psychologists and linguists a fairly clear idea of the existence of links between certain types of words and how these nodes combine to form a ‘word-web’ (cf. Aitchison 1994: 83). Several patterns emerge from these types of experiments which show the relative strength of certain lexical relations. If, for example, the stimulus has a clear antonym or is a member of a pair (e.g. *salt* and *pepper*), a very large majority of the subjects will respond with the opposite or second member of the pair. These results have been found to be very consistent across different experiments (cf. also EAT) and this evidence goes some way to establishing the structure of mental storage.[4]

Antonymy is considered by some (e.g. Princeton WordNet group, Gross et al. 1989, Miller 1990, Deese 1964 and 1965) to be a key structuring mechanism of lexical knowledge and is generally assumed to have a central position in at least the organisation of the adjectival mental lexicon (cf. 2.4.1). However, the question whether antonym pairs are stored as fixed combinations due to frequent co-occurrence and subsequent entrenchment or whether they are construed ‘on-line’

[4] There is a very large body of research on what priming experiments indicate about the structure of lexical storage which will not be discussed here in detail (cf. Harley 2008 for an overview).

as and when required has been much debated. In 2.4.2.2 evidence was presented from studies such as that of Charles et al. (1994), which investigated the difference between synonyms, direct antonyms and indirect antonyms to determine whether the relation of opposition was a conceptual or an associative relation. The results showed clear differences between synonyms and both types of antonyms but did not show clear evidence in support of one or the other scenario. Recent research by Paradis et al. (2009) suggests that there may be a certain number of highly canonical pairs which are stored as opposite pairs and are based on a strong associative relation, whereas the remaining members of the category of opposites are based on a conceptual relation and are construed when necessary by a feature comparison. This approach highlights the same division between canonical and non-canonical antonyms which has been pointed out a number of times before (among others Lyons 1977, Cruse 1986, Gross et al. 1989, Charles et al. 1994, Murphy 2003 and Paradis et al. 2007) and more recently also in Mohammad et al. (2008) who argue, from the perspective of natural language processing, that the large number of contrasting pairs which are not considered antonyms (in the Lyonsean sense) remain largely undocumented. They develop a distributional hypothesis which states that 'the degree of antonymy between two contrasting concepts […] is directly proportional to the distributional closeness of the two concepts' (ibid: 987). This ties in to the breadth of co-occurrence approach taken by Paradis et al. 2007 (cf. 2.5.3.2) and will be further investigated with the help of the data obtained from the behavioural studies reported below. The most recent research on antonymy, and the only other study to have employed several empirical methods, (Paradis et al. 2009) will be discussed in detail in Chapter 6 (cf. 6.3.1) in the light of the new evidence from the various studies in the present research.

A recent neurolinguistic study conducted in Korea (Jeon et al. 2009) has investigated processing of synonyms and antonyms contrastively using functional magnetic resonance imaging (fMRI) techniques in combination with a behavioural elicitation task. Jeon et al. (2009) advocate a feature-matching process and explain the spatial difference in the processing of synonymy and antonymy which can be seen in the fMRI data in terms of the difference between synonymy and antonymy of 'reversing the semantic meaning in one dimension and finding opposite features in case of antonyms' (Jeon et al. 2009: 453). The similarities in activation are, according to Jeon et al (2009: 455), due to the activation of the bundle of shared semantic features of the two concepts, whether antonymous or synonymous.

Based on experimental findings, Mohammad et al. (2008: 982), among many others, state that native speakers 'intuitively recognize different **degrees of**

**antonymy** [sic]'. This capacity, which is well supported by evidence, was already used in the present research in the GOE-rating task but the ability to discriminate between antonym pairs and to decide whether a pair of words is antonymous at all is a feature exploited by the design of the behavioural tasks below (in analogy to Gross et al. 1989 and Charles et al. 1994).

On the basis of the theoretical approaches and concepts outlined above, the following questions, which are similar to those discussed in 3.4, shall be addressed in the discussion of results following the presentation of the data:

(a) Which factors influence antonymic strength most and does this differ for different types of opposites (e.g. gradable antonyms, converses etc.)?

(b) Is there interaction between the individual factors?

(c) Do certain factors allow us to predict whether an antonym pair will be on a certain part of the antonymicity continuum, if such a continuum exists?

(d) Is there any evidence that speakers distinguish between types of opposition as proposed by Lyons (1977), Cruse (1986) and Cruse & Togia (1995)?

(e) Are we able to distinguish the influence of associative strength on reaction times from that of antonymic strength?

The difference between the data in the previous section and the data presented below is that this methodological approach does not allow time for the same amount of conscious reasoning which was possible in the judgement task. This enables an observation of the more automatic processes which govern antonym judgements and may result in the uncovering of differences in weighting of the individual factors which seem to influence these judgements.

Furthermore, the two questions below, which concern more general issues related to antonymy and lexical storage as well as linking into the idea of an antonym continuum raised in some of the questions above, will be touched upon but will only be discussed fully in Chapter 6 when the evidence from all three methodological approaches will be used to provide satisfactory answers.

(f) Does the data provide any evidence that there is a clear division between 'direct' and 'indirect' antonyms (as suggested by, for example, Gross et al. 1989)?

(g) Does the data show support for a continuum approach to antonymy?

## 4.2 Design of the antonymy decision task

The present research makes use of a task which is similar to a lexical decision task but involves antonym – non-antonym judgements instead of word – non-word judgements. This visual decision task consists of 120 word pairs which are presented visually to the subjects after a 10-pair trial run. Both words are presented at the same time in the following format: WORD1-WORD2 (e.g. HOT-COLD) in capital letters (font SYSTEM; size 36) in white on a black background in the centre of the screen. Each of the pairs is displayed for 1000ms with a break of 2050ms in between pairs. Since the task only contained 120 items (60 test and 60 control pairs – see 4.1.1.1 below), it was presented without any breaks as the total time of the task did not exceed eight minutes.

The programme used to present the visual stimuli was *Splice*, a signal processing programme[5] run on a MacPro ( MacOS 10.5.8) and, for both English experiments,[6] presented to the subject on ten individual monitors (ViewSonic 17'' PerfectFlat™ CRT Monitor E70f). The reaction times were recorded with purpose-built reaction time hardware (Reetz & Kleinmann 2003). The subjects all wore headphones (Sennheiser PX200) despite the fact that the task was purely visual to prevent being distracted by other participants' responses. All responses were given via custom-built two-button (Yes/No) boxes.

The task is structured akin to a lexical decision task. However, the decision the participants were asked to make was not between words and non-words but rather between antonym pairs and non-antonym pairs. Therefore, the task requires a meta-judgement which involves more conscious processing than is common in regular lexical decision tasks since it requires processing of the semantic content of both words as well as a decision about the relationship of the words which is not always solely based on morphological relatedness. Due to the more complex nature of the decision to be made and the number of letters in some pairs, the length of time for which the words were displayed on the screen had to be extended in order to allow sufficient time for the subjects to process the stimuli.

Not only do subjects have to retrieve both lexemes from long-term memory, they furthermore have to make a decision on the relatedness of the two words. The stimuli are thus carefully selected to mirror the word (antonym pair), pseudoword

[5] Developed by Henning Reetz (2008)
[6] In the German experiments, the stimuli were presented via a projector to all subjects at once. This is explained in further detail in 5.4.1.

(bad antonym pair or related but non-antonymic pair) and non-word (completely unrelated word pair) dimensions of a lexical decision task.

### 4.2.1 Stimulus selection

The selection of the word pairs used in this task was primarily guided by the results of the GOE-rating questionnaire discussed in the previous chapter. Those pairs which displayed particularly interesting discrepancies between associative and antonymic stength measures, for example *male:female* and its surrounding pairs, were chosen since closer investigation of said discrepancies might provide further insights into the factors which determine antonymic strength.

Other pairs chosen were those gradable pairs in which the criteria proposed by Cruse (1986) were most evident, for example the pairs which are distributed along the temperature scale. Converses made up a sizeable proportion of the 60 test pairs since they remain an under-researched sub-category of lexical opposition and the division between nominal and verbal converses seen in the questionnaire data (cf. 3.4.2) could prove to be an interesting point of study.

One last group of pairs which was included was that of words which have two different possible antonyms, for example *happy:sad* and *happy:unhappy* or *good:bad* and *good:evil*. These pairs are investigated more closely in the second experiment discussed in this chapter (cf. 4.4.2) but it was necessary to include them in this first task to provide a baseline against which the results of the second task can then be compared.

Every one of the lexemes in the 60 test pairs is matched to a control lexeme which matches the test word in frequency, word class, number of phonemes, syllables and letters as well as imageability.[7] The pairs which result from combinations of these control words are 50% completely unrelated (e.g. *subject:town*), 25% associatively related (e.g. *heavy:full*) and 25% synonymous (e.g. *pale:light*) to allow for different degrees of relatedness between the control words. On the one hand, this will help to determine whether antonyms are in fact recognised and judged at a different speed from unrelated items and, furthermore, whether any distinction can be made between the judgement of antonyms and synonyms and other associatively related word pairs. On the other hand, since the main focus of the study is that of the differences within the category of lexical

[7] Matched control items as well as imageability data were taken from the MRC Psycholinguistic Database (Wilson 1988).

opposition, the emphasis on the control items is somewhat less prominent. A list of all test items and their control pairs, including the relevant data, can be found in Appendix 4.

Morphological relatedness presents a problem in the design of the task since the fact that some pairs are morphologically related could lead to an immediate judgement which is based on morphology alone rather than semantic content. An attempt to avoid (or at least account for) this issue has been made by including pairs among the control items which consist of pairs in which one lexeme has a negative prefix but is not the opposite of the other lexeme (e.g. *competent:inconsistent*). The judgements of these cases will be instrumental in determining how much antonym judgements (especially within a constrained timeframe) are dependent on the morphological similarity of the two items or, indeed, on the presence of a negative prefix and on repetition priming, which is by all accounts an extremely effective form of priming.

To facilitate statistical analysis, all pairs were coded for the factors discussed in Chapter 3 (symmetry, gradability, morphological relatedness, word class, EAT score, GOE-rating score and t-score) as well as antonym type. The frequency-of-co-occurrence measures were converted into three measures: **high** (t-score of 10.0 or higher), **medium** (t-score between 2.0 and 10.0) and **low** (t-score below 2.0) and the GOE-rating scores were divided into the four groups discussed in 3.1.4 (**excellent** (1.0-1.79), **good** (1.8-2.5), **medium** (2.5-4.0) and **bad** (4.0-7.0)). The non-antonym pairs which were used as control items were divided into synonyms, associatively related pairs and non-related pairs to allow an investigation of the differences of various degrees of relatedness displayed by the pairs in this study. For a secondary analysis, the GOE-rating and t-score data was coded in continuous form to allow wider scope for statistical analysis.

Two groups were tested since each antonym pair had to appear once in its preferred and once in its dispreferred sequence. Both Task 1A and Task 1B contain a similar number of preferred and dispreferred sequences which were then simply reversed for the other task without altering the overall order of the word pairs.

### 4.2.2 Task and Procedure

The participants of the study received detailed information about the type of task before the experiment and were asked, upon having read the information as well as the instructions for the task (cf. Appendix 5), to sign a consent form which included some basic personal data (age, sex, level of education, handedness, variety of English, presence of any reading impairment (e.g. dyslexia)). Once the participants were seated at the individual computer stations in the behavioural laboratory, the instructions, which the participants had already read through before entering the room, were given again by the experimenter and the participants were given the opportunity to ask questions. The participants were asked to use the thumbs of each hand to press the **yes** and **no** buttons on their button boxes. All subjects used their preferred hand for **yes**. They were further instructed to wear the headphones to minimise disruption through other subjects' responses. The instructions to the subjects were loosely scripted so they would seem natural but so it could nevertheless be guaranteed that every participant was given the same instructions in the same manner.

The stimuli were presented via an iMac with *Splice* (see above) on the participants' individual screens. The group size varied from 2 to 14 during the pilot study of the experiment and was more consistent (2-6) in the second experiment.

The participants were first presented with a trial of ten items which consisted of five test items of varying antonymic strength as well as five control items of differing degrees of relatedness. Preceding the trial was the string 'X-X' which was used to indicate where on the screen the pairs would appear and to focus the participants' attention. This trial was run once and repeated if requested by the participants or deemed necessary by the experimenter. After the trial, a last opportunity was given to the participants to ask for clarification on the procedure of the task.

### 4.2.3 Participants

Since each antonym pair had to be run twice, once in each possible sequence, the participants were split into two groups. All 38 participants were aged between 18 and 37 and were all current undergraduates or postgraduates at the University of Oxford. They were recruited through advertisement and all participants were native speakers of British or American English with no special knowledge in linguistics. The majority of informants were right-handed; however, there were a small number of left-handed subjects (7/32). This should, due to the use of the preferred hand for the yes-response for all participants, not have any influence on the reaction times of these subjects or on the analysis. The informants took part on a voluntary basis and were paid for their participation in the task.

Some subjects had to be excluded from the analysis since they produced data with a higher than acceptable (10%) error rate. The design of this experiment, especially the selection of stimuli, complicates the setting of an appropriate cut off point for error rates since some of the word pairs used as targets (e.g. *debit:credit, tea:coffee, chase:flee*) are not considered opposite pairs by all speakers. Therefore, two possible methods to eliminate subjects which did not perform as well as the others in this task were considered: firstly, simply raising the error rate to 15% to account for the higher number of possible errors for word pairs which were coded as antonyms in the data but are not necessarily considered opposites by all speakers of the language. Secondly, to exclude some items which generated extremely high error rates and long reaction times due to the fact that some speakers would not consider them opposites (e.g. *bull:cow*) or because they were mistakenly coded as non-antonyms (e.g. *pale:dark*). This would allow for a decision on the basis of a more homogenous data set which subjects were to be excluded. The decision was made to use the second method as this seemed to be a more reliable method of ascertaining which subjects had performed below an acceptable standard. Using this technique, 6 subjects (4 from group A and 2 from group B) had to be excluded, leaving 16 subjects in each group with error rates no higher than 7% overall. Some of the items which caused indecision among subjects due to low antonymic strength and were thus excluded from the main statistical analysis will be discussed in 4.3.4.

## 4.3 Discussion of results

The statistical analysis of the data collected in the experiments was carried out with JMP (version 5.0.1.2).[8] The presentation and analysis of the behavioural data will be conducted in several stages. For all single-factor analyses, subjects were randomised and targets nested under the variable to ensure the validity of the data.[9]

In the first stage, a series of one-way ANOVAs were run to determine the individual influence of each factor on the independent variable (meanRT) and the factors will be discussed in turn. This will include analyses for all targets together as well as for certain sub-groups to determine whether some factors only influence a certain group of opposites (e.g. only gradable antonyms, converses or opposite

[8] JMP is a freeware statistical analysis programme.
[9] RSquare is reported to indicate the explanatory power of the analysis. An RSquare greater than 0.30 is considered acceptable in the context of linguistic data (Henning Reetz, personal communication).

pairs with excellent GOE-rating). In some cases the analysis will also be split according to lexical categories.

In the second stage, several multi-variate ANOVAs will be run on the data to determine interaction between certain groups of criteria. Once again, these analyses will be carried out for the whole data set as well as for appropriate subsets to investigate differences according to antonymic strength. The third stage is less statistics-driven and will discuss some of the idiosyncrasies in the data by looking at either individual items or sub-groups which show particulary unusual patterns, especially when compared to the GOE-ratings discussed in the previous chapter. This will conclude the analysis of the data and is be followed by general discussion.

### 4.3.1 Single factor analyses

#### *4.3.1.1 GOE-rating*

The first of the individual factors to be analysed is the GOE-rating. The results of the judgement task were presented in Chapter 3 and therefore this section will only deal with the relation in which these results stand to those of the behavioural experiment. To illustrate the distribution of the data, the questionnaire results were split into the four groups proposed in 3.1.4, excellent (1.0 – 1.79), good (1.8 – 2.99), medium (3.0 – 4.99) and poor (5.0 – 7.0). These groups are solely based on the results of the judgement tasks and have not been changed for the statistical analysis in this chapter. Thus, there are some idiosyncrasies which stem from the GOE-data which result in some unusual distributions of the independent variable in the discussion below.

The first graph below (Figure 4.2) shows the distribution of MeanRT for all pairs from all four groups which were selected as stimuli for the behavioural task (for a full list, please see Appendix 4). The ANOVA showed that the distribution below is significant ($F (2, 1385) = 14.2788$; $p \leq 0.0001$; RSquare: 0.58). A further t-test ($F (2, 1385) = 58.930$, $p \leq 0.0001$) shows that the differences between all three conditions in the analysis are significant. The least square means[10] are as follows: **excellent** (752.184), **good** (817.675) and **medium** (849.722). There were not enough instances of pairs in the **poor** category of GOE-rating to warrant a statistical analysis, which was thus excluded from all analyses.

[10] In the discussion of the results of all ANOVAs in this study, meanRT refers to the least square means since those are more accurate than the mean reaction times since they take into account variation between subjects and targets. However, for the qualitative analyses carried out in the second half of this chapter, mean reaction times are used.

This is an expected distribution since the GOE-score was assumed to be a stable indicator of antonymic strength. The pairs which were surprising in their results in the behavioural task compared to their GOE-rating score are ones differing along the gender divide (e.g. *king:queen*). These pairs will be discussed later in the investigation of the influence of associative strength on the behavioural results since they display reaction times which do not correspond to their rating in the judgement task.

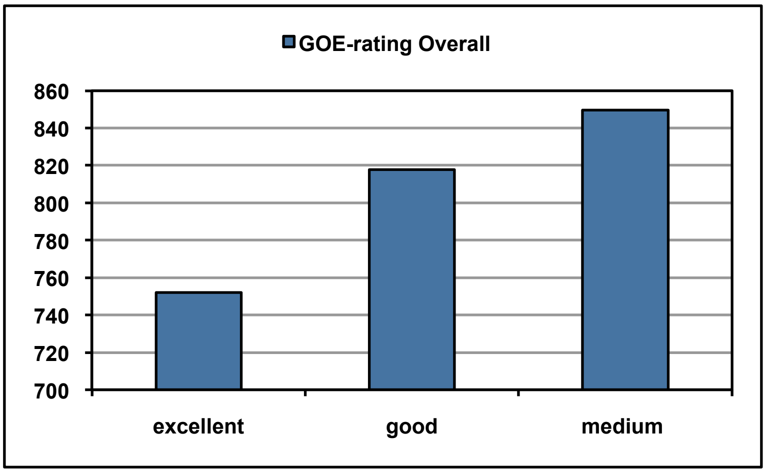
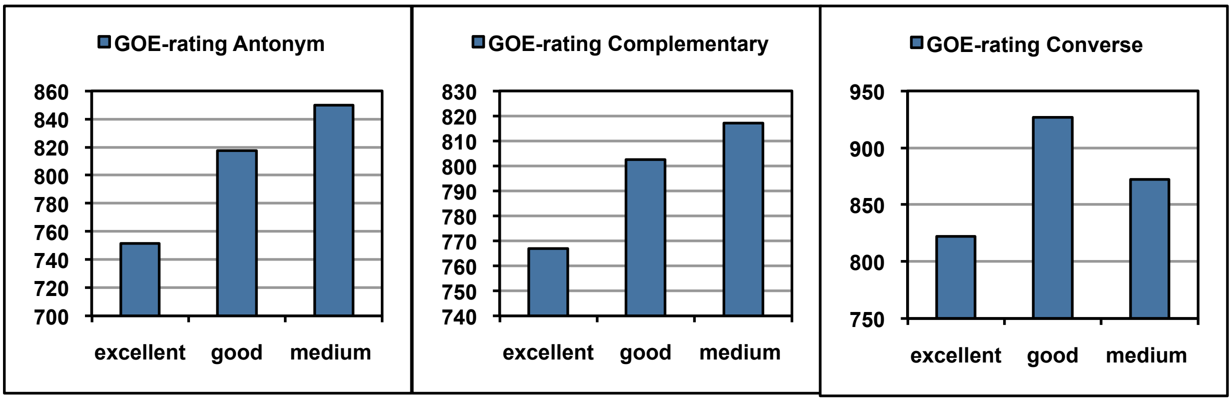



**Figure 4.2** *MeanRT (in ms) by GOE-rating*

Figure 4.3 below shows MeanRT by GOE-rating split according to antonym type. The three main antonym types investigated in the present work – antonymy (in the Lyonsean sense), complementarity and converseness - will be looked at separately to see in which cases GOE-rating ties in most closely with the behavioural results. The hypothesis will be addressed that the congruence of the two measures should be greatest in the antonymy and complementary categories and weaker in the converse category.




**4.3a** **4.3b** **4.3c**

**Figure 4.3 (a-c)** *MeanRT (in ms) by GOE-rating and antonym type*

Figure 4.3a shows meanRT divided into GOE-rating groups (as above) for all gradable antonyms included in the study. The ANOVA for antonyms only (F (2, 807) = 39.0473; p ≤ .0001; RSquare: 0.60) indicates that excellent (751.507) antonyms are recognised faster than good ones (817.213) which in turn are recognised faster than medium ones (1075.383). The meanRTs of all three categories are significantly different from each other.

In the case of the complementaries, the distribution is not as clear-cut. The ANOVA (F (2, 374) = 5.502; p ≤ 0.0044) still displays a strong degree of overall significance of the results but the significant differences between the GOE-rating categories are not as easily definable as in 4.3a. Figure 4.3b shows the results in meanRT for all complementaries included in the study. While the category of antonymy contained mainly adjectival pairs, the complementaries contain a wider range of lexical categories. A t-test with contrast shows that there is a significant difference between **excellent** and **good** complementaries (766.840 and 802.525 respectively), and between **excellent** and **medium** (816.990) pairs. There is, however, no significant difference between **good** and **medium** pairs. The reasons for this distribution and their implications will be discussed below in 4.3.4.

The last figure above (Figure 4.3c) shows the distribution of converses. This is an interesting case since there is only a very limited number of converses which are classed as **excellent** or even **good** antonyms and none of the converses included in this study are adjectival. The ANOVA for all converse pairs only (F (2, 262) = 10.2312; p ≤ 0.0001; RSquare: 0.51) shows strong statistical significance between **excellent** and **good** (811.285 and 926.762 respectively) and **excellent** and **medium** (872.145). **Good** and **medium** are not significantly different. The reason for the faster reaction times in the medium category may be explained by the results for the pairs around the base pair *male:female,* which displayed a strikingly different pattern from that which was apparent in the GOE-rating and will be discussed in more detail in 4.3.4.

### *4.3.1.2 Frequency of co-occurrence*

The influence that frequent co-occurence and subsequent entrenchment of certain lexical items has on the strength of their relationship and the reliability of their connection in the mental lexicon is indisputable. However, the question is whether associative strength is the key factor in antonym judgements or whether the degree of opposition between the two concepts involved is more crucial to the relationship. This is a question which, despite contributions by numerous researchers in the field, has not been answered satisfactorily to date. Behavioural tasks which require the participant to make as quick a decision as possible, such as the ones conducted in the present study, may be able to give certain clues which lead towards a clearer picture of the situation since participants are required to rely on the fastest mechanisms possible. In this case we may see an effect which prioritises items which have strong associative connections but are not generally judged as strongly antonymic in judgement tasks. If that should be the case, which measure is to be taken as more indicative of the antonymic strength of a pair of lexical opposite remains to be discussed (cf. 4.3 & 6.2.2).

The statistical analysis of degree of frequency of co-occurrence was carried out in three stages: firstly for the overall dataset, secondly divided by antonym type and lastly by GOE-rating score. All three analyses will be considered in turn and then compared.

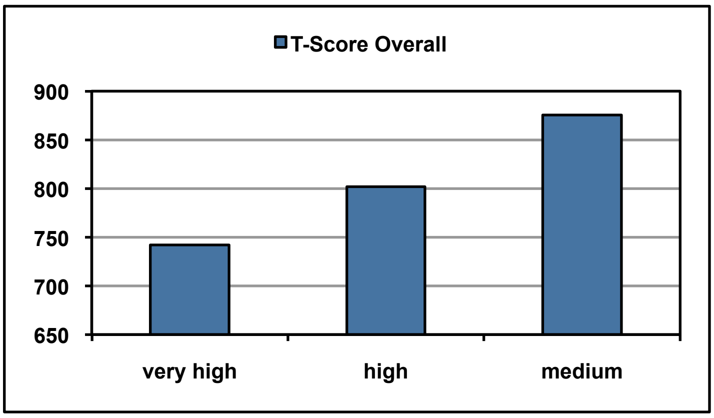
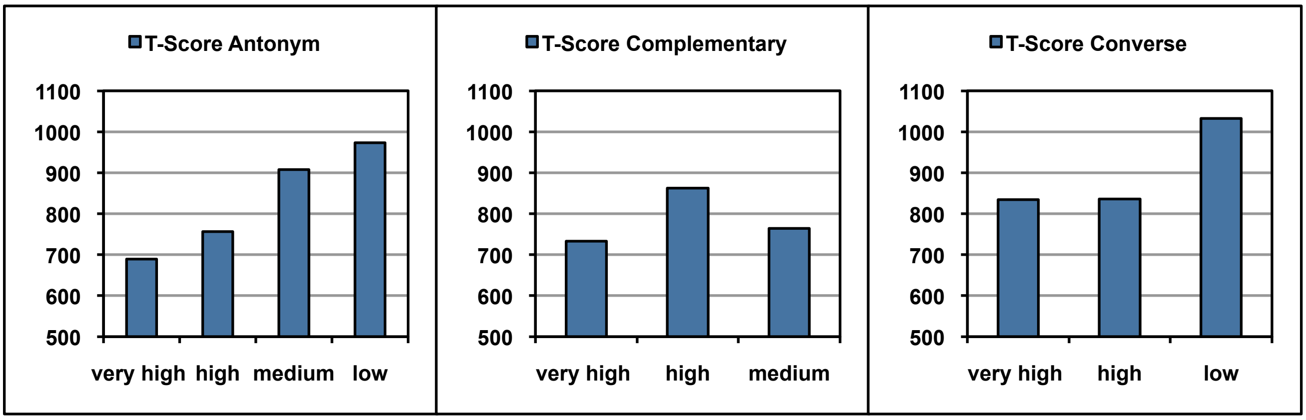



**Figure 4.4** *MeanRT (in ms) by t-score*

In Figure 4.4 we can see that the results of the first ANOVA (F (2, 1385) = 33.5398, $p \leq 0.0001$, RSquare: 0.58) show a significant difference between the categories **very high** (742.062) and **high** (801.869) and **medium** (875.541). There were not enough instances of items with low t-score to warrant a statistical analysis. It needs to be pointed out, however, that a t-score in the medium category (between 2.0 and 10.0) is still considered evidence of co-occurrence greater than chance would predict whereas if a pair has a t-score in the category low (< 2.0), the co-occurrence rate of the two items can be attributed to chance. The fact that co-occurrence seems to be a factor which influences reaction times to a significant degree in the overall analysis of the data is not surprising since it has been proven by extensive corpus studies that antonyms do indeed occur more frequently in the same context than chance would predict (cf. Jones 2002, Jones & Murphy 2005 and Jones et al. 2007).

In Figure 4.5, the ANOVA carried out in the three datasets which only contained antonyms of a certain type is illustrated. The three charts show very different distributions of frequency of co-occurrence. This is the first, and only, time that the three different antonym types included in this experiment display markedly different patterns.




**4.5a** **4.5b** **4.5c**

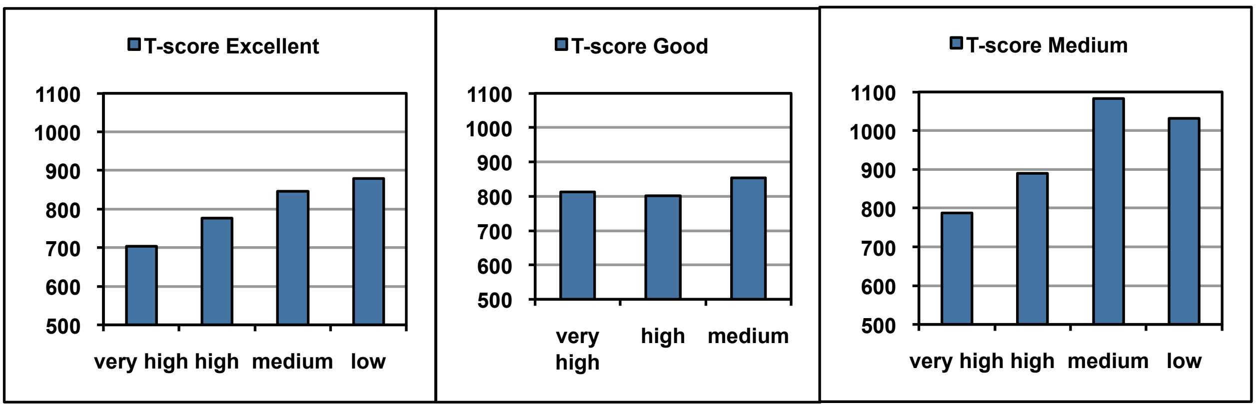
**Figure 4.5** *MeanRT (in ms) by t-score and antonym type*

Figure 4.5a shows results reflecting a ‘classic’ distribution which shows that the analysis ($F (3, 807) = 126.1151$; $p \leq .0001$; RSquare: 0.60) displays significant differences between all four conditions: **very high** (690.051), **high** (756.527), **medium** (907.553) and **low** (972.966). Since the group of **antonyms** contains many of the central gradable adjectival pairs as well as the members of the clusters around them, this is not altogether surprising.

The middle figure (4.5b) shows the different conditions in the group of complementaries and while the analysis ($F (3, 400) = 34.8620$; $p \leq .0001$; RSquare: 0.56) also results in overall statistically significant differences between some of the conditions, the distribution is less intuitive than that in the previous chart. The top two categories are significantly different from each other (**very high**: 732.838; **high**: 862.193) and the lowest category is significantly different from both **very high** and **high** (684.500). The meanRT of pairs with a medium t-score is not significantly different from that for pairs with a very high score (764.793). This result had to be re-evaluated after post-hoc analysis, since there were only a very limited number of stimuli in the condition **low** for complementaries and thus the results are not representative and can thus be considered irrelevant for the discussion below (cf. 4.3.3). The other result of note is that those complementaries which obtained a high t-score display a significantly slower mean reaction time than those which obtained medium t-scores, while a very high t-score resulted in similar reaction times to the medium pairs. Here, the question to what extent the exact value of the t-score (or the absolute, rather than relative, frequency of co-occurrence) is crucial in antonym (or other relational) judgements must be asked. This will be discussed further in 6.1.1.5 and in the discussion of the experimental results in this chapter (cf. 4.3.5).

Figure 4.5c shows the conditions for the group of converses only. There are no converses in this analysis which obtained a medium t-score and thus the chart only shows three conditions. The overall analysis is statistically significant (F (2, 262) = 9.9612; p ≤ 0.0001) but there is barely any difference between the conditions **very high** (833.772) and **high** (836.345). Pairs in both of these conditions are recognised significantly faster than those in the condition **low** (1031.6977). From the data displayed in this chart, it appears that the degree of strength of the association does not have a direct influence on these less conventionalised pairs. However, the fact that there is a co-occurrence rate which is higher than chance does seem to influence the reaction time considerably.

The last set of charts in Figure 4.6 displays the results of the criterion of frequency of co-occurrence divided by GOE-rating. There were not many instances of the GOE-rating **poor** in the overall questionnaire which is why the analysis was only carried out for the three categories shown below. For Figure 4.6a, the results of the analysis of the four conditions were carried out on a dataset which only contains pairs which received excellent GOE-ratings in the judgement task. The analysis (F (3, 857) = 77.4039; p ≤ 0.0001; RSquare: 0.57) shows that the differences in reaction times between the conditions are highly significant. However, there is no significant difference between the **medium** (846.620) and **low** (878.725) conditions while **very high** (702.979) and **high** (775.901) are both significantly different from each other and from the bottom two conditions. This is, again, an expected distribution since opposite pairs rated excellent do occur more frequently in context. It is, however, interesting that the GOE-rating category **excellent** contains pairs from all four conditions (e.g. *even:uneven* is in the category **low**).




**4.6a** **4.6b** **4.6c**

**Figure 4.6** *MeanRT (in ms) by t-score and GOE-rating*

The analysis of the pairs in the GOE-rating category **good** (Figure 4.6b) displays fewer differences between reaction times and thus less congruence between the two measures (GOE-rating and t-score). The ANOVA shows no overall significance (F (2, 414) = 5.5700; p < 0.0041; RSquare: 0.55) and there is no significant difference between the **very high** and **high** conditions (813.445 and 801.230 respectively) and only the condition **high** differs significantly from **medium** (853.438). The fact that there is no difference between the higher categories leads us to the same question as the analysis of Figure 4.5c above. However, the differences between **very high** and **high** on the one hand and **medium** on the other are not as marked and the fact remains that **medium**, as already mentioned above, is nevertheless indicative of a frequency of co-occurrence which is greater than chance.

The last figure (4.6c) shows the distribution of the conditions in the dataset for pairs which obtained a medium GOE-rating. There are still a significant number of pairs in this category which obtained high or very high t-scores which, in itself, creates doubts as to how useful an indicator of antonymic strength frequency of co-occurrence is. The ANOVA (F (3, 173) = 17.5160; p ≤ 0.0001; RSquare: 0.63) showed significant differences between three groups of conditions: **very high** (787.49) and **high** (890.322) differ significantly and both are significantly different from **medium** (1083.314) and **low** (1032.19). The fact that the condition medium (t-score), once again, results in slower reaction times and whether this can be traced back to an error in the coding of the conditions will be discussed in 4.3.6 where some of the methodological issues encountered in this study will be discussed. Overall, it is clear that while t-score does have an influence on meanRT, antonymic strength also plays an important role since the GOE-rating groups display progressively slower reaction times. The fast times for very high (t-score) in this category result from the very fast times obtained by a small number of pairs. These are discussed in more detail in 4.3.4.1. Before the impact of these analyses on the theoretical basis of the experiment is discussed (cf. 4.3.5), the other factors which may influence antonymic strength are evaluated in turn.

*4.3.1.3 Sequence*

The question whether the sequence in which antonyms are presented has an influence on how easily the pair is recognised as antonymic has been under discussion in recent work on antonymy. Paradis et al. (2007) found that there was no overall difference in the judgement task they carried out between antonym pairs which were displayed in their preferred and dispreferred sequence[11]. In the current research, both the questionnaire (cf. 3.2) and the behavioural task (cf. 4.2) were designed to detect any potential influence sequencing has on antonym canonicity. The preferred and dispreferred labels were given on the basis of the co-occurrence data extracted from the BNC. All pairs which showed a considerable discrepancy in co-occurrence in the two sequences were labelled preferred/dispreferred in the coding for the experiment whereas all pairs which did not show any such preference or for which no co-occurrence data could be obtained were labelled ‘no preference’. The pairs which received this label were not considered in the analysis carried out below.

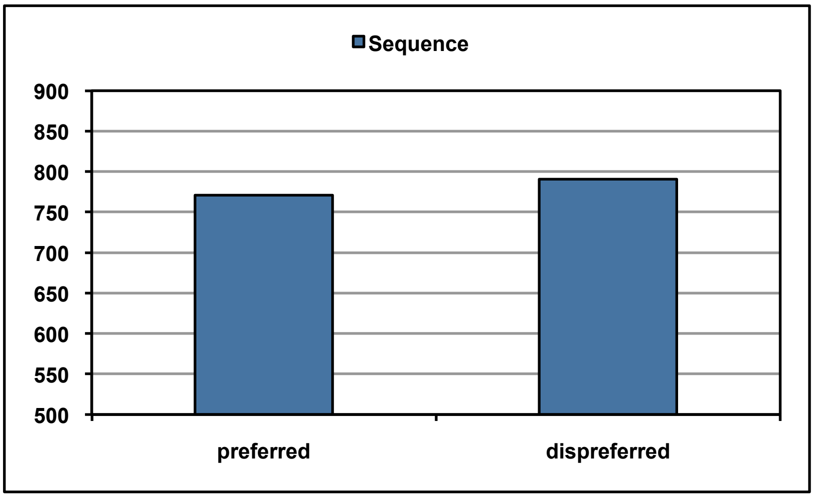



**Figure 4.7** *MeanRT (in ms) by sequence*

As Figure 4.7 above shows, there is barely any significant difference in reaction time between antonym pairs occurring in their preferred and dispreferred order when the groups are considered as a whole (F (1,1347) = 3.866; p < 0.0495; RSquare: 0.56). In a qualitative post-hoc analysis, however, there are some pairs which do show large discrepancies between the W1/W2 and W2/W1 combinations which account for the slight significance in the results of the ANOVA. Figure 4.8 below shows both combinations and their corresponding mean reaction times for each pair. The dark blue line represents the reaction times given for the pair in the order in which they appear in the legend whereas the light blue line stands for those times which resulted from the inverted combination.

It is evident from the two graphs in the figure below that there are some in which the order of presentation does have an influence on a pair’s perceived antonymic strength.

[11] For this purpose, preferred and dispreferred sequence was determined by the BNC co-occurrence data.

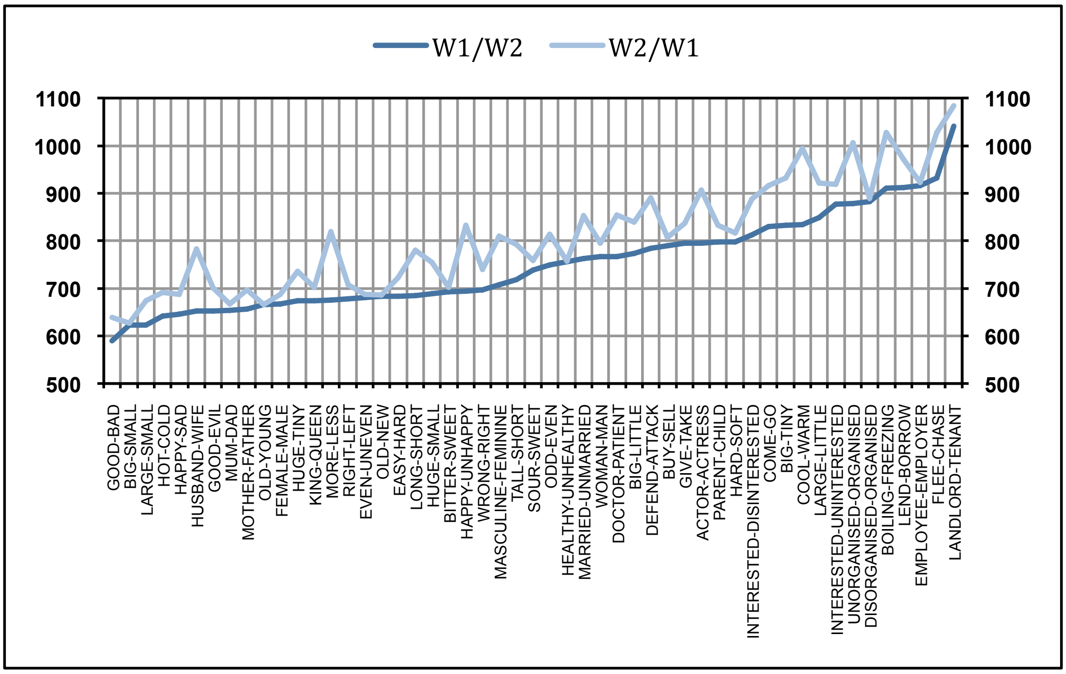



**Figure 4.8** *MeanRT (in ms) by opposite pair and sequence*

Overall, the results of the behavioural experiments support the criteria put forward for sequencing (e.g. by Jones 2002 – cf. 2.2.2) since those pairs which do display a preference do so within the parameters set by previous scholars of this phenomenon (e.g. positive before negative, more before less). A comparison of the pairs which display differences in reaction time and those pairs which showed large discrepancies in the GOE-rating will be carried out in 6.1.2.3.

*4.3.1.4 Symmetry*

Since symmetry is a criterion which only applies in cases of gradable opposites, it is not surprising that an overall analysis of the complete dataset does not show symmetry to be a statistically significant factor as far as its influence on reaction time is concerned. However, when only the subset of gradable antonyms are analysed, the results look very different. Figure 4.9 shows the results of the oneway ANOVA carried out on gradable antonyms only which indicate that symmetry does influence the speed with which a decision is made about the antonymicity of the pairs in question ($F\ (1, 632) = 29.0858$; $p \leq 0.0001$; RSquare: 0.59).

This finding ties in with the analysis of the results of the judgement task in 3.4.1, which showed very clearly that pairs which are equidistant from the midpoint of the scale along which they are distributed are judged to be better examples of antonymy than those which are asymmetrically distributed. Since this only applies to a subgroup of antonymy, namely those pairs which are representatives of a property which can be constructed as a scale, it was expected only to have an influence in this particular group. Another interesting result of the statistical analysis is that the asymmetrical pairs displayed a standard error which was over twice as high as that of symmetrical pairs (20.577 vs. 8.0744).

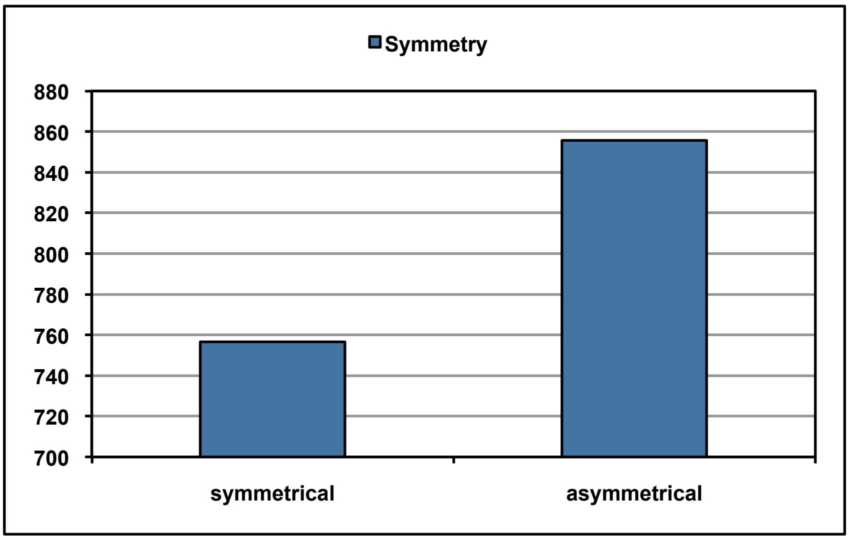



**Figure 4.9** *MeanRT (in ms) by symmetry*

In the analysis of the German judgement task and behavioural data, undertaken in Chapter 5, symmetry will receive a little more attention since, at least in the questionnaire data, it seems that it is a more important factor in German than it is in English. Whether this can be substantiated by the behavioural data will be discussed in 5.4.3.

*4.3.1.5 Morphological relatedness*

The criterion of morphological relatedness has been much discussed. There is no disagreement about the fact that the 'best' antonyms, those which are judged to posess the greatest degree of antonymic strength, are morphologically unrelated (cf. among many others Cruse 1986 and Paradis et al. 2007). However, while unrelated opposites are spread along the whole canonicity continuum, morphologically related pairs are only to be found in the top third of the continuum. It should therefore be reflected in the data that both the most and the least strongly related opposite pairs are morphologically unrelated while the related pairs are usually considered good opposites with much greater consistency.

One problem which was encountered in the comparison of morphologically related and unrelated items in the behavioural task was that of item length. As discussed above, item length has a considerable influence on reading time and thus reaction time and, if not treated carefully, will lead to a misinterpretation of the results. When the oneway ANOVA was carried out on the complete dataset of correct responses, it showed a significant result for morphological relatedness (F (1, 1301) = 61.8392; p ≤ 0.0001; RSquare: 0.57) with the related items being around 60ms slower than the unrelated ones (837.468 vs. 770.632).

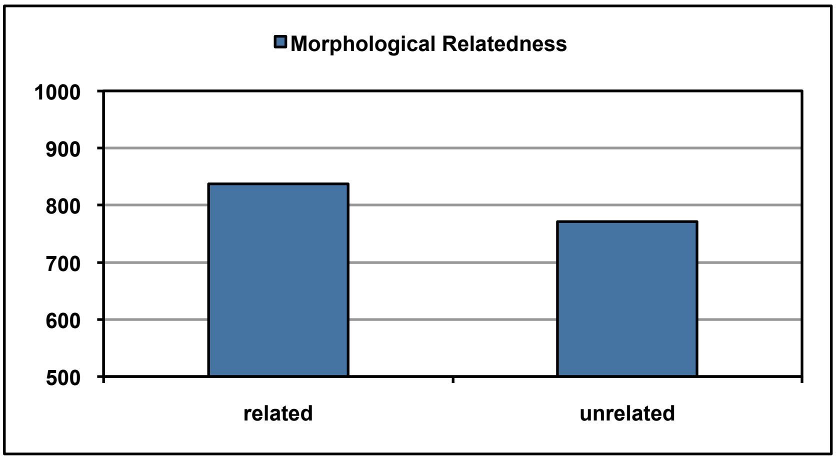



**Figure 4.10** *MeanRT (in ms) by morphological relatedness*

The results of a second oneway ANOVA, carried out on the same dataset with all items over 18 characters in length taken out, no longer show significance. In this second analysis, there is no significant difference between the two conditions (F (1, 1185) = 2.1737; p ≤ 0.1406; RSquare: 0.56). The mean reaction times for the unrelated and related conditions are 770.54 and 786.585 respectively. Bearing in mind that some long items still remain in this dataset and that a number of them are morphologically related, this presumably still influences the results to a certain degree. However, it illustrates what the GOE-rating results in the previous chapter already showed and emphasises the fact that the antonymic strength of morphologically related items is generally much more consistent than that of unrelated items which, at the top end of the scale, may rely on associative strength in addition to their already considerable antonymic strength and, at the bottom end, require additional processing. This, as well as the reasons for this distribution, will be discussed from a theoretical point of view in 6.1.2.2 and 6.2.2.

#### *4.3.1.6 Word class*

The question of what role, if any, word class plays in antonym judgements and whether the fact that most good antonym pairs are adjectives can shed light on some of the processes and criteria involved in judgements and processing of antonymy has already been discussed in 3.4.3 and 3.4.4 and will be considered again in 6.1.3.1 as a criterion of good opposition as well as in connection with the kind of processing involved in antonymy recognition (cf. 6.2.2).

The largest number of pairs in both the judgement task and the behavioural experiments were, of course, adjectival. However, in both tasks a significant number of verbal and nominal pairs were included for purposes of comparison. Many of these pairs, both nominal and verbal, are converses, which led to two separate analyses being carried out – one in which the dataset included converses and one in which it did not. The outcome of both analyses was remarkably similar and thus only the results of the analysis of the complete dataset are presented below (Figure 4.11).

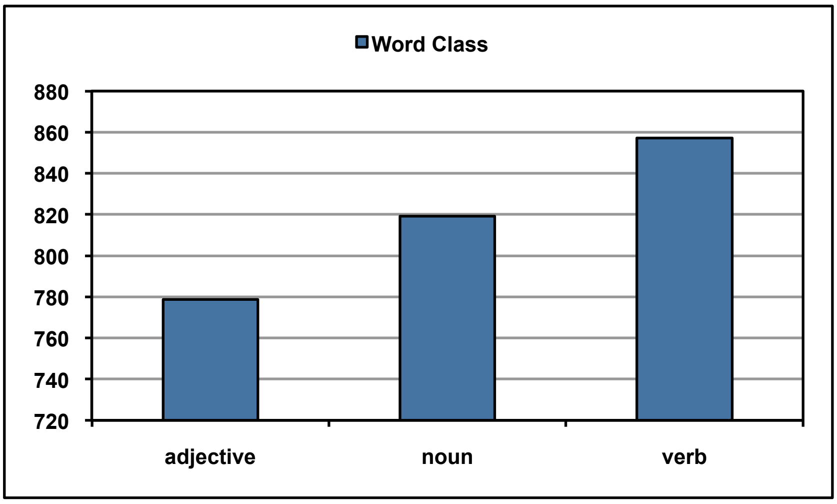



**Figure 4.11** *MeanRT (in ms) by word class*

The figure above shows the results of a oneway ANOVA, which are strongly significant ($F(2, 1500) = 32.3821$; $p \leq 0.0001$; RSquare: 0.59). All three groups are significantly different from each other and are spaced very evenly. The mean reaction time for adjectives (778.579) is roughly 35ms faster than that for nouns (819.080) which, in turn, is roughly 35ms faster than that for verbs (857.247). What is striking is that these results do not match the data obtained in the judgement task very well since there the verbal converses at least showed much more variability and generally higher antonymic strength than their nominal counterparts. This, however, may be due to the fact that some nominal pairs (cf. gender pairs below in 4.3.5.1) obtained much faster reaction times than their GOE-rating score would predict. Therefore, the analysis was carried out a second time without the four pairs (*mum:dad*, *mother:father, man:woman* and *king:queen*) to which the following criteria apply:

a) very high t-score
b) bad match of meanRT and GOE-rating
c) very strong associative strength as two members of a pair

The results for the second ANOVA are still strongly significant ($F(2, 1396) = 52.0008$; $p \leq 0.0001$; RSquare: 0.5941) but the three conditions are no longer significantly different from each other. Figure 4.12 below shows the 'new' distribution of results. The adjectival pairs (778.581) are still recognised significantly more quickly than the nominal and verbal pairs (865.696 and 857.175 respectively). This seems to be a very robust result which, of course, is partly determined by the fact that such a large number of the **good** and **excellent** pairs in the judgement task are adjectival.

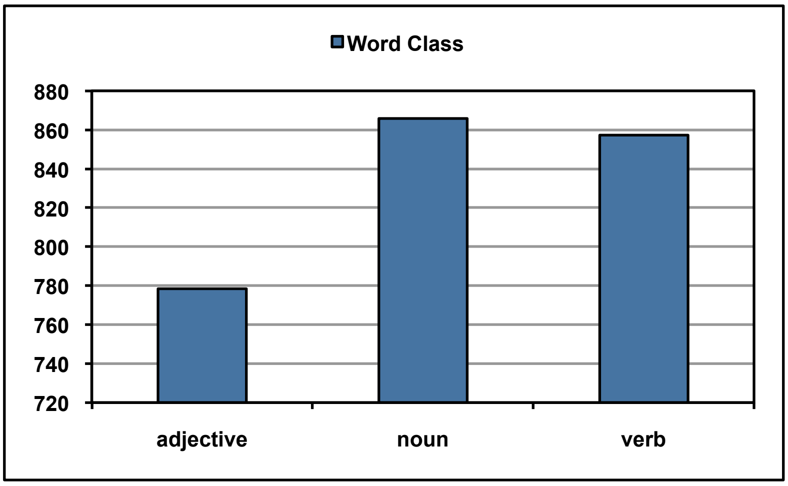



**Figure 4.12** *MeanRT (in ms) by word class (amended)*

The difference between Figures 4.11 and 4.12 can be seen in the mean reaction time for verbal and nominal pairs. While in the first analysis, nominal pairs were recognised significantly faster than verbal ones, in this second ANOVA, there is no significant difference between the two conditions. There is even a tendency in the opposite direction since the verbal pairs are, on average, recognised about 25ms more quickly than the nominal ones. Thus the explanation of different degrees of complexity in internal category structure given in 3.4.2 and 3.4.3 may well be applicable to these results as well.

*4.3.1.7 Antonym type*

The distribution of meanRT by antonym type in Figure 4.13 below shows an unsurprising result. The two categories of antonym (783.135) and complementary (784.444) are almost identical in mean reaction time whereas the reaction times to converses were significantly slower (849.174). The oneway ANOVA confirms that the differences shown in the figure are statistically significant ($F(2, 1529) = 29.7352$; $p \leq 0.0001$; 0.58).

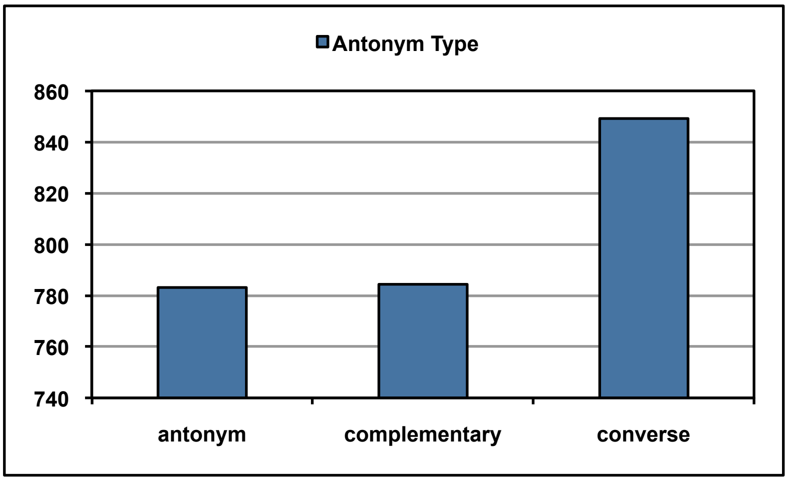



**Figure 4.13** *MeanRT (in ms) by antonym type*

The difference between the two more conventional categories on one side and converses on the other is one which has appeared time and again in the previous sections in various contexts. The reason why converses constitute such a different class is most likely partly because the antonymic relationship of the two members of a converse pair is considerably more complex than that in other opposites. Some of the criteria which seem to be crucial to good antonymy are not, or only latently, present in some converse pairs. As far as the experimental results are concerned, one factor which may have played a role in the differences in results of the quantitative analysis is the fact that the other two antonym types outnumber the converses in both the judgement task and the behavioural experiments. However, when the data is considered from a qualitative perspective (cf. 3.4.2 for judgement task data) in 4.3.5.1 below, it becomes clear that there is more at play than a simple design fault in the experimental setup.

### 4.3.2 Interaction of criteria

To determine how the individual criteria analysed above interact with each other, this section considers some of the factors discussed in the previous section to investigate how they influence each other and how their combination influences mean reaction time. Some of the variables above only apply to subsets of the data (e.g. symmetry, gradability and morphological relatedness) and are therefore not included in the analysis in this section. Antonym type, word class, t-score and GOE-rating are all factors which will be investigated here. The analyses are complicated slightly since not all factors are evenly distributed and thus the statistical treatment of these cases is difficult.

The first analysis carried out was an ANOVA investigating the interaction between GOE-rating and t-score. Only the GOE-rating categories **excellent** and **good** were included to see if the t-score plays any role in the distribution of pairs between these two categories. Low t-scores were also excluded since there were not enough instances of these to warrant an analysis. The ANOVA is significant for both t-score ($F (2, 1192) = 39.3111; p \leq 0.0001$) and GOE-rating ($F (1, 1192) = 40.1072; p \leq 0.0001$) as well as for the two factors crossed ($F (2, 1192) = 17.9676; p \leq 0.0001$).

The analysis shows a significant interaction between t-score and GOE-rating. It can be seen that the associative strength of the pairs, represented by the t-score, seems to have a greater effect in the behavioural task than it does in the GOE-rating. However, it does not solely determine the distribution shown in the analysis since pairs in the combination **very high/good** are not significantly different from the pairs with medium t-scores. There are some unexpected distributions in the data which show that the interaction between the two features is not straightforward and is modulated by additional features.

In terms of antonym type, the interaction with GOE-rating provided significant results which support the classical distribution of antonym types. A t-test, which provides difference values for the individual conditions, showed that gradable pairs with excellent GOE-rating were the ones which were recognized fastest overall while complementaries and converses are clustered together with more similar reaction times. T-score also has an effect here, which added to the antonymic strength of the pairs with very high t-scores. As expected, the fastest times are achieved by the pairs which have very high t-scores, excellent GOE-rating and are adjectival gradable opposite pairs. However, the distribution below this initial top level is not influenced by any one of these factors alone but by different factors to different degrees. Associative strength will have to be considered separately from antonymic strength as it adds another dimension which does not absolutely correlate with the measures of antonymic strength.

### 4.3.3 General discussion

The previous sections have given an overview over the data, first by taking a close look at each of the criteria which have been proposed to influence antonym judgements and then by looking at the interaction of these criteria. It has become clear that not all criteria have the same amount of influence on antonymic strength and that some only apply in certain cases. The three criteria which did not result in any significant differences were morphological relatedness, sequence and symmetry. Of these three, symmetry is a criterion which only applies to antonym pairs which are gradable along a certain scale and, when examined only in this group, symmetry proved to be strongly significant.

Overall significance was seen in the remaining four criteria: antonym type, word class, t-score and GOE-rating. The first two only show significant differences between one condition on the one hand and two on the other while, in the case of t-score and GOE-rating, the situation is not quite as simple (as can be seen from the additional analyses in 4.3.2 and 4.3.5). Before moving on to the discussion of idiosyncratic data (4.3.5), the data discussed above will be compared to the data gathered by corpus analysis and in the judgement task. For this purpose only the opposite pairs on the SIZE continuum are considered since they are a particularly interesting cluster in English due to the fact that they are arranged around two (or more?) canonical pairs at the centre of the cluster.

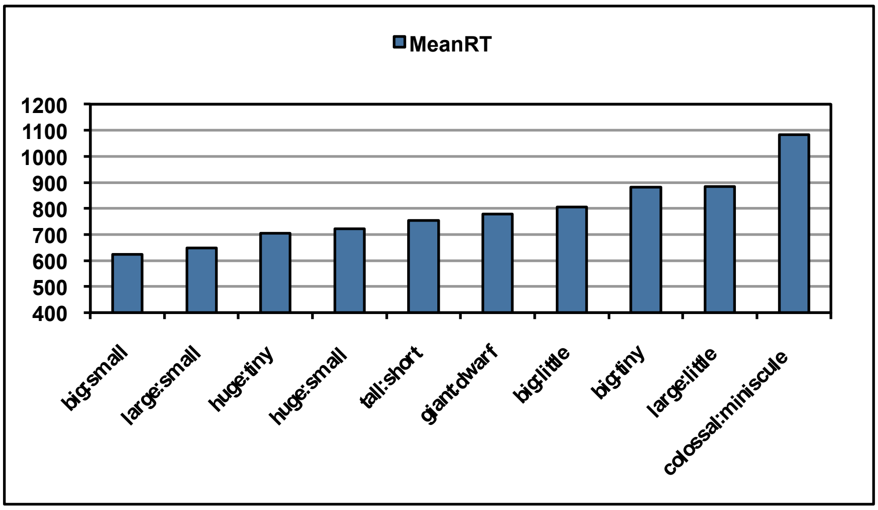



**Figure 4.14** *MeanRT (in ms) for pairs on the SIZE continuum*

Figure 4.14 shows the reaction times for all pairs on the SIZE continuum. The distribution of reaction times is relatively linear with no clear break between the central pairs and the more peripheral ones. In fact, the two central pairs, taking into account reaction time data only, are not the two pairings which are generally (cf. K.J. Miller 1998) taken to be the canonical centre of the size adjectives: *big:little* and *large:small*. One pair which may have to be considered with a caveat only is *colossal:miniscule* since its item length considerably exceeds that of all the other pairs in this group: this will have had an effect on recognition speed and therefore also reaction time. The analysis does not draw any conclusions from the low reaction times obtained by this particular pair.

Given that symmetry did have an effect on gradable adjectives, it is surprising to see how far to the left *huge:small* appears in the chart above and that it has a faster reaction time than some symmetrical pairs, most notably *big:little*. This, unlike the fast reaction times of the nominal pairs mentioned in 4.3.1.6, cannot be attributed to an unusually high t-score and thus considerable associative strength. Both Figure 4.15 and Table 4.1 show that neither the t-score of *huge:small* (1.53) nor its GOE-rating (2.25) seem to indicate a particularly high antonymic strength.

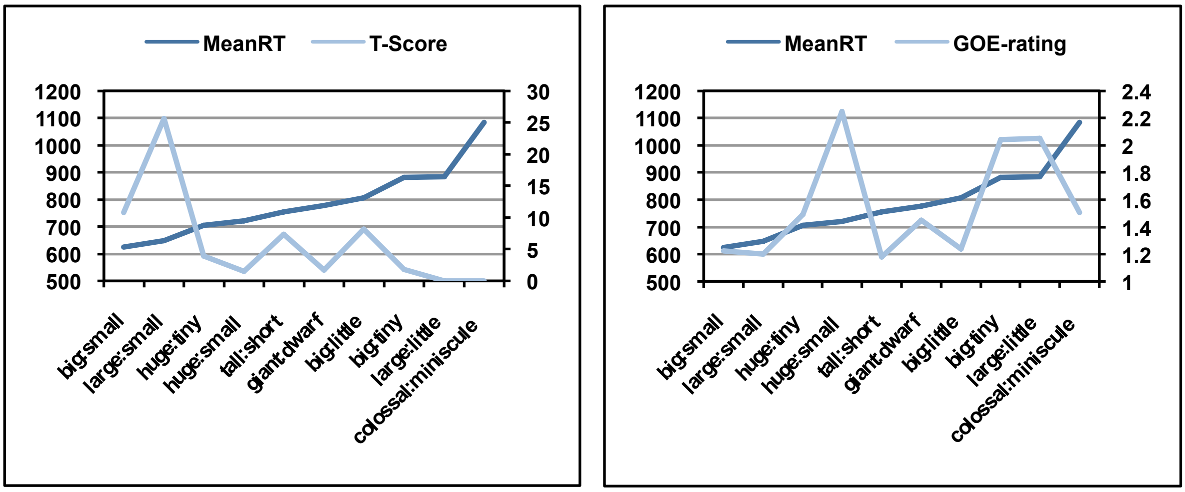



**Figure 4.15** *Comparison of results for* SIZE *adjectives (meanRT in ms)*

However, from the results of the whole dataset, it can be seen that both t-score and GOE-rating seem to influence meanRT in some way. Unlike in an 'offline' task, in the behavioural experiments, several processes are at work in antonym recognition since the decision of whether a pair is antonymic or not has to be made under very different conditions.

A close connection between two lexemes can have a number of effects in a task which relies strongly on automatic (or seemingly automatic) processes (cf. 4.1); the strong relationship can either have a facilitatory function (which semantic and associative priming make use of (cf. 4.4.2.2 – size adjectives)) or an inhibitory function which slows down the reaction of a participant to a particular prime-target combination. Although the task in Experiment 1 was not designed as a priming task, a certain amount of priming is inevitable since a small delay before reading the second word is to be expected.

The central pairs form a curious cluster of their own since the connections between the four main size adjectives, *big*, *large*, *small* and *little*, do not pattern exactly in the way many speakers of the language would have predicted. Even though the pairs *big:little* and *big:small* have almost identical results in the two other empirical measures, their reaction times differ significantly which leads to the hypothesis of an inhibitory effect in the case of *big:little* where the participants expect *small* rather than *little* as partner for *big*. This is discussed further in the analysis of the data from Experiment 2 which was designed to specifically investigate these differences (cf. 4.4.2.2).

**Table 4.1** *Comparison of results for SIZE adjectives (by meanRT)*

| | MEANRT | T-SCORE | GOE-RATING |
|---|---|---|---|
| big:small | 624.69 | 10.79 | 1.225 |
| large:small | 648.285 | 25.67 | 1.2 |
| huge:tiny | 705.015 | 3.9 | 1.49 |
| huge:small | 721.345 | 1.53 | 2.25 |
| tall:short | 755.065 | 7.38 | 1.18 |
| giant:dwarf | 777.489 | 1.72 | 1.45 |
| big:little | 806.41 | 8.13 | 1.24 |
| big:tiny | 882.23 | 1.82 | 2.04 |
| large:little | 884.625 | 0 | 2.05 |
| colossal:miniscule | 1083.765 | 0 | 1.505 |

When it is taken into account that the conditions for this kind of methodology differ considerably, the results seem to highlight the same criteria as important in the judgement of lexical opposition as the results discussed in the previous chapter. Some of the factors, for example the difference in antonymic strength for different word classes, have been supported by the data discussed in this section. Others, for example frequency of co-occurrence, have proven to be a more significant factor in this type of task than in the judgement task. However, this does not answer the question whether frequency of co-occurrence determines and enhances antonymic strength or whether it is a result of the strong antonymic relation of the two lexemes involved. T-score results will be discussed further in the following section and the 'chicken-and-egg' question of the interplay between frequency of co-occurrence and antonymic strength will be addressed in 6.1.1.5 and 6.2.3.

### 4.3.4 A closer look – idiosyncratic data

Following on from the qualitative analysis of the two gradable antonym clusters discussed in the previous section, which displayed results only differing very slightly from those which could have been predicted by the data gathered from the corpus analysis and the judgement task, some other antonym pairs will now be considered in greater detail. As previously mentioned, there were several instances in the data where the discrepancies between the results of the other experimental measures (GOE-rating and t-score) have provided starting points for discussions (cf. for example 3.4.2 and 3.4.3). This section will consider two of these cases in which the results of the behavioural experiment do not match those obtained in the judgement task, which is usually considered a reliable measure of antonymic strength (cf. among others Murphy and Andrews 1993, Paradis et al. 2009): the cluster of pairs along the gender divide which has already shown interesting differences in t-score and GOE-rating, and the group of converses included in the experiment. Both of these cases were already discussed in the previous chapter (3.4.2 & 3.4.3) and therefore the discussion will continue where Chapter 3 left off and add evidence provided by the behavioural data.

#### *4.3.4.1 Another look at* GENDER

There are some peculiarities among the gender pairs which have already been mentioned in the discussion of the individual criteria of word class and converses. Some of the nominal pairs which, surprisingly, scored very low on the GOE-rating task (e.g. *man:woman* and *mother:father*) received very fast reaction times in the behavioural experiment. In this section, these particular pairs are examined in greater detail and an explanation of the idiosyncrasies displayed by the different sets of data is sought.

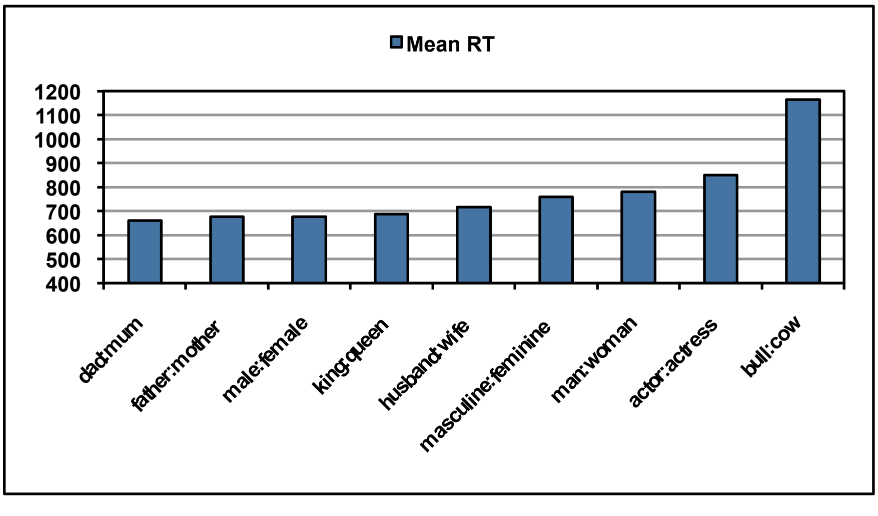
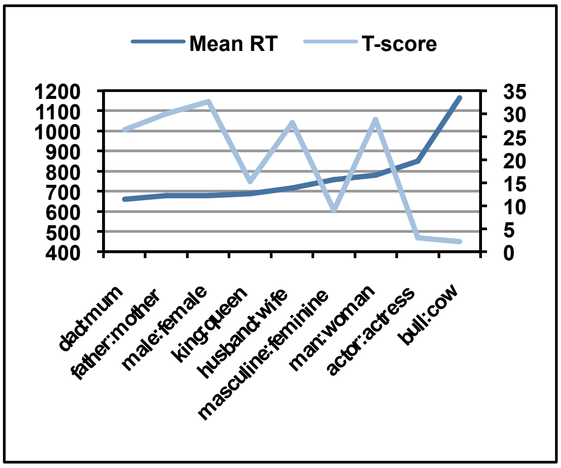


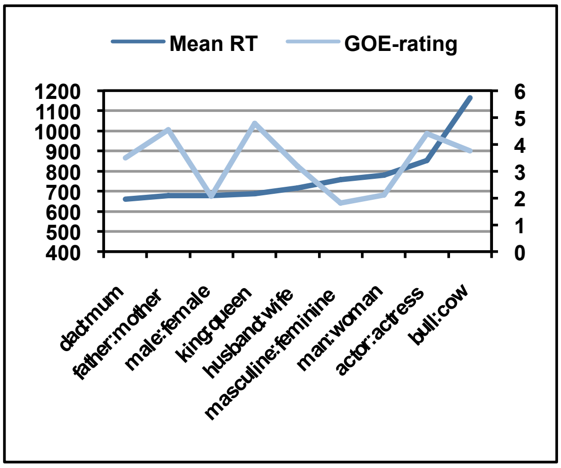

**Figure 4.16** *MeanRT (in ms) for* GENDER *pairs*

Figure 4.16 above shows the mean reaction times of all gender pairs included in the experiment. It is evident, even at first glance, that these results are markedly different from those discussed in 3.4.3 where not even the two adjectival pairs had a score high enough to put them in the **excellent** category of the GOE-rating (*male:female* 2.06 and *masculine:feminine* 1.81). The two pairs which result in the fastest reaction times, *dad:mum* and *father:mother*, were ones which were both rated in the **medium** category of the GOE-rating. *Bull:cow* is only included in the results in 4.16 for completeness' sake since there are not enough responses for this pair (5 out of 32 in total) to draw any conclusions from its individual analysis.

The chart below (Figure 4.17) attempts to show the relationships between the three independent measures used in this study to illustrate the differences between their results. It is clear from the graphs, as well as from the numbers in Table 4.2, that there is no immediately obvious explanation. However, there are several points of interest which, under close investigation could prove important from a theoretical as well as a methodological perspective.




**4.17a**




**4.17b**

**Figure 4.17** *Comparison of results for* GENDER *pairs (meanRT in ms)*

Figure 4.17a shows the comparison between meanRT and t-score for the gender pairs while Figure 4.17b shows a comparison of meanRT and GOE-rating. In 4.17a, the greater the difference between the two lines, the better the two measures match since high t-scores are a better indication of associative strength whereas lower meanRT corresponds to a stronger relationship between the two lexemes. In 4.17b, the closer the lines are together, the better the match between meanRT and GOE-rating score since in both those cases, a lower score is an indication of a stronger relation.

Once again, the graphs are slightly misleading since, at first glance, they give the impression that there is no congruence between the measures at all. One observation which must be made is that the (high) t-scores of some of the word pairs in this group (cf. also Table 4.2 below) seem to have much more of an influence on the reaction time than on the GOE-rating (see discussion in 3.4.3). This seems to be the case especially at the extremes of the spectrum since the distances are greatest at the low and high end of the RT continuum in Figure 4.17a. Figure 4.17b shows an even more varied picture and the GOE-ratings seem to have very little to do with the results of the lexical decision task. The three pairs with the highest scores, the two adjectival pairs mentioned above and *man:woman*, seem to show some correspondence between the scores but no firm conclusions can be drawn from this.

**Table 4.2** *Comparison of results for GENDER pairs (by meanRT)*

| | MEAN RT | T-SCORE | GOE-RATING |
|---|---|---|---|
| dad:mum | 660.410 | 26.47 | 3.5 |
| father:mother | 676.855 | 29.99 | 4.55 |
| male:female | 677.370 | 32.54 | 2.06 |
| king:queen | 688.465 | 15.23 | 4.775 |
| husband:wife | 717.710 | 28.01 | 3.175 |
| masculine:feminine | 758.477 | 8.94 | 1.81 |
| man:woman | 780.780 | 28.67 | 2.11 |
| actor:actress | 851.225 | 2.96 | 4.4 |
| bull:cow | 1165.085 | 2.18 | 3.755 |

Also taking into account the direct comparsion of the figures in Table 4.2, the data seems to confirm that the t-score plays a more important role in the behavioural task than it does in the self-paced ‘offline’ judgement task. It seems to be the case that the extremely high associative strength of the items with the unusually high meanRT, *dad:mum*, *mother:father*, *king:queen* and, to a lesser extent, *man:woman*, aids processing and thus also speeds up the antonym decision. It may even be the case that the ‘normal’ process of antonym recognition presented above (cf. 4.3.3) is superseded by associative strength in cases like the those presented above where the two members of the pair are very much a unit and co-occur together so frequently as two sides of a coin (rather than two polar opposites) so they still display enough antonymic strength in their association to be classed as ‘opposites’ in this online lexical decision task since the decisions for these items were not only made quickly but subjects produced very few errors when reacting to those pairs.

These items are, however, not considered canonical antonyms and are therefore not part of the small group of excellent opposites which, in addition to displaying high antonymic strength, also benefit from the entrenchment of their lexical relation. Therefore, this could be an effect generated by the design of the experiment and a clue that antonym judgement may be conducted in stages and is facilitated by high associative strength since the connection between the two lexemes, albeit not necessarily an antonymic one in the first instance, is activated much more quickly than in pairs whith lower rates of co-occurrence and thus the decision of whether a certain pair is an opposite pair can also be made more quickly. One way to determine whether this hypothesis is at least in part correct may be through a neurolinguistic investigation of these processes, which would involve re-running the lexical decision task as an EEG study to be able to analyse the ERP data which could provide very interesting results. At present, however, an investigation of the behaviour of the converse pairs may provide additional support for this hypothesis.

*4.3.4.2 Converses – a different kind of opposite?*

The GOE-rating results for converses displayed an interesting distribution of scores within and between the two word classes represented in the group of converses in this study. In 4.3.1.6, an ANOVA carried out to determine whether there was any difference in the results of the behavioural study for the three different opposite types showed that while complementaries and antonyms were extremely similar in meanRT, converses overall took much longer to recognise and react to. This applied equally to verbal and nominal converses.

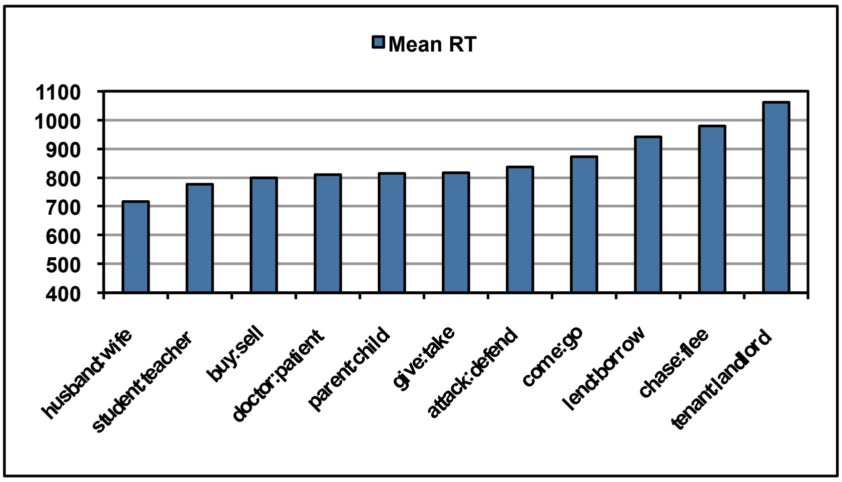
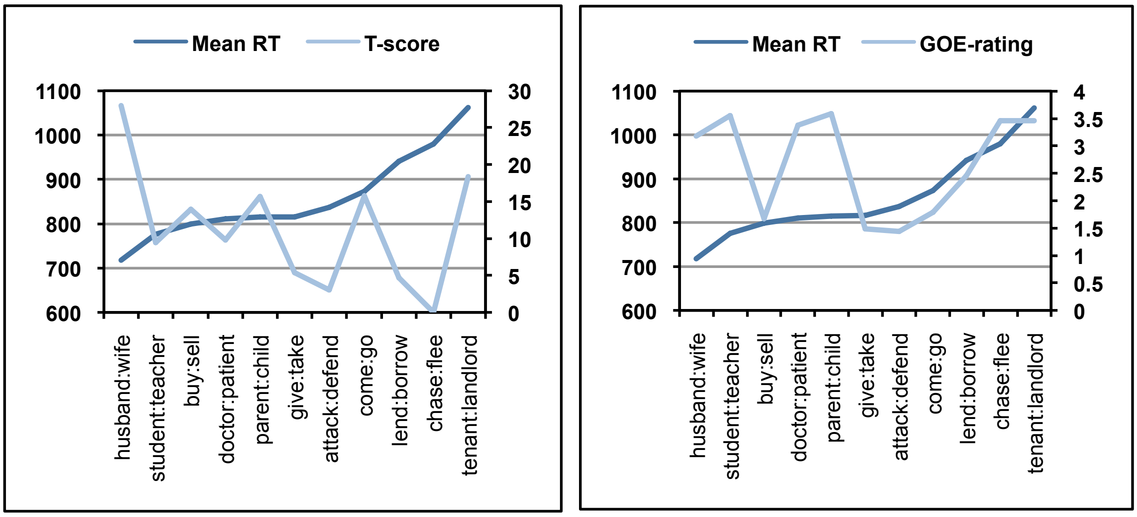



**Figure 4.18** *MeanRT (in ms) for verbal and nominal converses*

In the GOE-rating data a clear distinction between nominal and verbal converses could be seen, and this was initially attributed to the more complex nature of the nominal concepts and thus the greater difficulty of construing the binary scales necessary for antonym construal. The mean reaction times for each converse pair in Figure 4.18 show no such distinction between the verbal and nominal pairs. It seems that, despite the fact that many of the verbal pairs received better ratings in the judgement task, those pairs obtain slower reaction times.




**4.19a** **4.19b**

**Figure 4.19** *Comparison of results for converses (meanRT in ms)*

As in the case of the results for the GENDER pairs analysed above, the two charts in Figure 4.19 directly compare meanRT and t-score (4.19a) and meanRT and GOE-rating (4.19b). In terms of a match of GOE-rating and reaction time, the verbal converse pairs seem reasonably well matched whereas the nominal pairs show greater divergence. In part, this can once again be attributed to the phenomenon observed in the previous section (cf. 4.3.4.1): nominal pairs with very high t-scores obtain much faster reaction times than their GOE-rating would predict.

**Table 4.3** *Comparison of results for converses (by meanRT)*

| | MEAN RT | T-SCORE | GOE-RATING |
|---|---|---|---|
| husband:wife | 717.71 | 28.01 | 3.175 |
| student:teacher | 775.98 | 9.46 | 3.555 |
| buy:sell | 798.78 | 13.99 | 1.67 |
| doctor:patient | 810.42 | 9.76 | 3.375 |
| parent:child | 814.87 | 15.66 | 3.59 |
| give:take | 815.885 | 5.356 | 1.49 |
| attack:defend | 837.1 | 3.07 | 1.44 |
| come:go | 872.666 | 15.74 | 1.785 |
| lend:borrow | 941.395 | 4.67 | 2.45 |
| chase:flee | 979.83 | 0 | 3.46 |
| tenant:landlord | 1062.185 | 18.37 | 3.465 |

The high t-score but corresponding slow reaction time of *tenant:landlord* may be due to a large sample of legal documents or adverts for housing in the BNC, which inflate the t-score to the extent that it ceases to be an accurate measure of associative strength. This hypothesis, if true, would explain this single instance without questioning the overall conclusion that the weighting of the criteria for good opposition is somewhat different in a judgement task from a behavioural experiment task where decisions about antonymicity are made under severe time constraints.

### 4.3.5 Summary

The combined results of the single-factor analyses as well as the interaction of the criteria and the conclusion we can draw from the data are discussed in this section. Seven individual factors were considered in the analyses above, and several have already been discussed extensively and will therefore only be mentioned as far as their previous discussion has an effect on the overall conclusions.

Two out of the seven factors, **morphological relatedness** and **sequence,** did not display a statistically significant effect on the independent variable in the analyses carried out above. Sequence has been seen to have a considerable effect only on certain pairs whereas a large number of pairs do not display any preference of a particular sequence. This criterion has been seen to have no overall effect (cf. also Paradis et al. 2009). Morphological relatedness also does not significantly influence the decision in the lexical decision task but will be discussed further once the results of the German data have been presented since the German evidence is not as clear-cut as the English data presented above (cf. 5.1.1 and 5.4.5.2).

The results of the single-factor analysis for **symmetry** tie in very well with the GOE-rating task which confirmed that equidistance from the midpoint of a scale (Cruse 1986) is an important criterion in determining antonymic strength. Symmetrical opposite pairs score better than their asymmetrical counterparts even when the latter contain one of the members of the base pair (cf. 3.4.1) while the symmetrical pairs do not. Symmetry did not emerge as a significant criterion overall but when only the gradable pairs in the study were considered, the analysis showed that symmetrical pairs obtained significantly faster reaction times than their asymmetrical counterparts despite often being very similar in t-score.

The analysis of **antonym type** showed that there is a difference in reaction and processing time between antonyms and complementaries on the one hand and converses on the other hand. The fact that converses take longer to process could be due to the more complex relation between the two members of a converse pair – a thought already put forward above – but the fact that all converses in this study are either nominal or verbal while the bulk of the pairs in the remaining two groups is adjectival could also play a role. As the analysis of the criterion of **word class** showed, there is a significant difference in reaction time between the three lexical categories which seems to confirm the hypothesis put forward in 3.4.3 at least to a certain degree since, in line with the predictions of the hypothesis, adjectives are recognised faster than verbs which, in turn, are recognised faster than nouns.

The remaining two factors, **antonymic strength** (GOE-rating) and **associative strength** (t-score), form a slightly different category since both of them are already empirical measures related to every single pair. Both are shown to have a statistically significant influence on the outcome of the lexical decision tasks in Experiment 1. Overall, **antonymic strength** seems to have the expected effect with the pairs which received an **excellent** GOE-rating score providing the fastest reaction times. The picture which emerges in the analysis of **associative strength** initially seems very clear (cf. 4.3.1.2) but the results obtained in a more detailed analysis do not match those for antonymic strength, which gave cause for post-hoc investigations which provided interesting results: some of the pairs which had not been rated very highly were among the pairs which obtained the fastest reaction times. These pairs were mainly nominal (e.g. king:queen, husband:wife), many of them shared certain features and all of them had very high t-scores. This shows that the influence of frequency of co-occurrence and subsequent entrenchment of the connection in a pair is of greater importance in a task like the behavioural experiment than in an 'offline' rating task like the GOE-task.

It was expected that this factor would figure more prominently in the behavioural experiments than in the judgement tasks since the strong associative relation between two lexical items which co-occur frequently is an important tool which facilitates processing. It seems from the data that high associative strength makes a considerable difference also in those items which are not judged to be very strongly related in the GOE-rating task. It is clear that this high degree of relatedness aids recognition and speeds up the decision process – in some cases considerably (e.g. *mum:dad* or *man:woman*) – and does not simply aid the faster recognition of those pairs which are very strongly antonymically related. To what extent the exact value of the t-score (or the absolute rather than relative frequency of co-occurrence) influences antonym (or other relational) judgements is difficult to determine. A higher than chance co-occurrence (t-score above 2.0) seems to have an overall effect whereas the extreme facilitation seen in some of the examples mentioned above only occurs with an extremely high t-score (> 25). To investigate whether there are different processing aspects involved in the recognition of antonym pairs with varying degrees of associative strength, an ERP study would have to be conducted (cf. 7.3 & 7.4).

## 4.4 Experiment 2

The second behavioural task is of the same kind as the one described in 4.2.2 but the experimental design and the presentation of the word pairs differ slightly. This task was considerably shorter than the previous one and, rather than looking at the factors which influence antonym judgements in general, was designed to answer only one very specific question: do lexemes which have two possible opposites have a preferred antonym or is the antonymic strength for both approximately equal?

### 4.4.1 Methodology

To attempt to answer this question, the design of the antonym decision task used in Experiment 1 had to be modified. It was shortened significantly and the group of target stimuli consisted of only 13 word pairs (cf. Table 4.4).

#### *4.4.1.1 Stimulus selection and experimental design*

All word pairs from the questionnaire which had more than one possible opposite were selected for this task. There are three groups of stimuli: one where the base antonym forms one morphologically related and one unrelated pair, one where both pairs are morphologically related and one where both are unrelated.

**Table 4.4** *Target stimuli used in Experiment 2*

| TYPE | BASE | ANT 1 | ANT 2 |
|---|---|---|---|
| related-unrelated | happy | sad | unhappy |
| | healthy | sick | unhealthy |
| | even | odd | uneven |
| | married | single | unmarried |
| | | | |
| both related | organised | unorganised | disorganised |
| | interested | uninterested | disinterested |
| | | | |
| both unrelated | short | tall | long |
| | good | bad | evil |
| | hard | easy | soft |
| | right | left | wrong |
| | sweet | bitter | sour |
| | | | |
| quartet | little | big | large |
| | small | big | large |

The case of the size adjectives is, as mentioned before, a rather special one and has been included since for the two pairs, there are four different combinations and the question of which pairs with which has long been discussed (cf. 3.4.1.3). Since a task with only 13 pairs and their corresponding controls would be very short and would not give the participants the necessary time to get used to the procedure, filler items were inserted pseudo-randomly (for a complete list see Appendix 6) to increase the total number of stimuli to 50. The items used as fillers were pairs from Experiment 1 and the corresponding control pairs. To ensure a minimum level of antonymic strength throughout the experiment, only pairs from the good and excellent groups were used. Since each base antonym had to be presented once with antonym 1 and once with antonym 2, two almost identical tasks (Task 2A and Task 2B) were created in which the only difference is the second member of the 13 target pairs.

The experiment was designed as a priming experiment using delayed presentation of the second member of the pair. The first lexeme of the word pair was presented for 300ms, removed and 450ms later replaced with the complete word pair which was displayed for 800ms. Due to restrictions of the software used for the experiments, the 450ms delay between the offset of the prime and the onset of the stimulus could not be reduced any further without disrupting the recording of accurate response times. It was also not possible to leave the first member of the pair on the screen and to simply add the second member after 300ms. Instead, a fixed-width font was used to ensure that the first member of the pair was displayed in exactly the same place on the screen both on its own and in the word pair. This slight alteration in the originally planned design, however, is not expected to have affected the results significantly since the method used is very similar to a regular priming paradigm. The slightly longer priming time of 750ms in total may allow for a wider spread of activation and may allow for post-lexical processing (cf. Harley 2008). The time interval between the word pair and the next prime was 1450ms. The 50 word pairs were presented in five blocks of ten and each block was separated from the next by a 5200ms pause.

As before, the stimuli were presented on individual computer monitors in a behavioural laboratory with a maximum of eight subjects being tested at once. The words were presented in white on a black screen in capital letters in font size 36 in a fixed-width font (Monaco) to ensure the prime appeared in the same place as the identical first word of the target word pair. Participants used two-button boxes to record their responses. The task was presented using *Splice* on the same experimental equipment as described in 4.2 and the reaction time data was once again recorded by a custom-built data-collection device (Reetz & Kleinmann 2003).

#### *4.4.1.2 Task and procedure*

Before the start of the task, the participants were asked to read information on the task and then sign a consent form which also contained questions on relevant personal information (age, sex, level of education, handedness, variety of English, presence of any reading impairment (e.g. dyslexia)). Following this, they were given the instructions for the task and were asked to read them thoroughly. The experimenter then explained the procedure and answered questions. A test task of ten items, none of which was included in the actual task, was then run to ensure the participants' comprehension of the requirements of the task. The first three items of the experimental task were either fillers or control stimuli and were not included in the analysis to ensure that the participant had settled into the task before the relevant data was obtained.

As in the previous experiment, subjects were asked to wear headphones despite the task being exclusively visual to reduce external noise and prevent being distracted by other subjects' responses. They were further instructed to hold the response box like a games console and use their thumbs to press the buttons. All subjects used their dominant hand for the 'yes' response.

#### *4.4.1.3 Participants*

There were two groups of fourteen participants each. The only criteria for the task were that the participants were native speakers of English and not dyslexic. All participants were undergraduate or graduate students at the University of Oxford and took part in the study voluntarily but none had participated in Experiment 1. The average age of participants was 23 (age range: 19 - 32).

Since the task was extremely short, it was run in conjunction with another study (a cross-modal priming task). The participants were compensated for their time. The study has ethics approval from the University of Oxford's ethics committee (CUREC).

Unfortunately, four participants in each group had to be excluded due to high error rates on the target stimuli and therefore the data from only ten subjects in each group was used in the analysis below.

### 4.4.2 Analysis of priming data

Since in this task both members of the word pairs did not appear on the screen at the same time but one member of the pair was first presented alone, this was in effect a standard priming task with a relatively short stimulus offset-onset asynchrony of 450ms. This task was intended to provide evidence for whether certain lexemes which have two possible good opposites show a stronger association with one than with the other.

The hypothesis which underlies this task is that the amount of priming seen in the task should be greater the stronger the pairs are related and thus the results should show faster reaction times for pairs which are more strongly related. Furthermore, there may even be cases of inhibitory priming since the task design generates the expectation of an antonymically related pair and every one of the test items has two possible very good opposites which could be chosen. If one of these two choices is more closely related to the prime, it should be activated more easily and the presentation of the other, more unexpected, opposite partner should cause an increase in reaction time since it would require additional processing effort to discard the original, preferred lexeme (cf. 4.4.2.2; *big:little*).

Since antonyms of the canonical type (which all target items are), as discussed in the preceding chapters, is a relation which seems to depend both on association resulting from frequent co-occurrence and on a strong semantic relationship dependent on feature overlap between the members of an opposite pair, it is irrelevant in this case to determine whether priming is associative or semantic. The data will be presented in a short overview after which the different groups in Table 4.4 above will be discussed individually in greater detail.

#### *4.4.2.1 Overview*

The analysis of the data is carried out for 26 instead of 30 pairs since four of the pairs did not produce methodologically satisfactory results. The morphologically related pairs based on *interested* and *organised* resulted in too many errors as the display time was relatively short and the number of letters of these items, especially the derived antonyms, seems to exceed that which can be read by the participants in the time given. Thus, these pairs have been excluded from the analysis completely. Figure 4.20 below shows the mean reaction time and standard deviation for the remaining pairs.

Both measures are interesting to consider and give valuable clues as to the strength of the association between the two lexemes. It can not be determined whether the accelerated processing and thus reaction time to these items is indeed based on the strength of their antonymic relationship or simply a reflection of their associative strength. However, since the task required the participants to make a decision about the antonymic relatedness of the pairs and since all pairs investigated in this study have tested as excellent, canonical antonym pairs in the GOE-rating study it will be assumed that the priming which can be seen in the experimental data is a reflection of the degree of antonymy between the two items.

The mean reaction times overall vary between around 400-700ms. Some items, like *sweet:sour:bitter* which, unlike any of the other opposites, represent a true triad and have little overlap in meaning between all three items, will also have to be discounted in the detailed analysis since (in this particular case) more than half the reactions to the pair *sweet:bitter* were zero reactions, which skews the mean reaction time too much and does not leave enough data to draw solid conclusions.

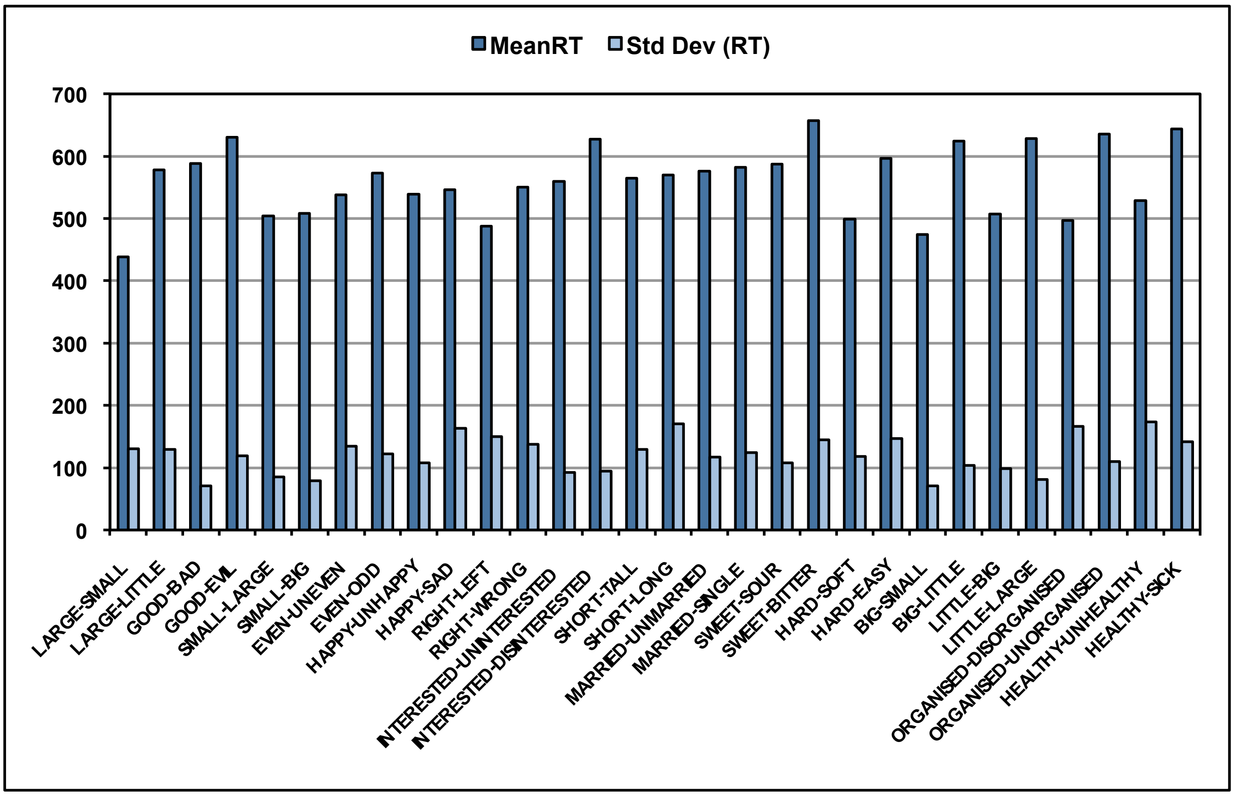



**Figure 4.20** *MeanRT and Std Dev (in ms) for test items in Experiment 2*[12]

The standard deviation is less than 180ms in all cases and in some it is significantly less. The relatively small number of participants in this experiment makes it somewhat difficult to draw conclusions from the standard deviations. In a large

[12] Pairs are presented according to base adjective (with each of the possible partners).

number of cases, the pair with the higher standard deviation is also that with the longer reaction time but these two features do not correlate strongly enough to indicate more than a tendency. The direct comparison between the pairs, especially in the case of the four size adjectives, does however, show some interesting results.

#### *4.4.2.2 Comparison of individual pairs*

The pairwise analysis is conducted in three parts according to the groups presented in Table 4.5. The first group to be analysed is that of the related-unrelated pairs, followed by the unrelated-unrelated pairs and then, before moving to the contrastive analysis in the next chapter, the quartet of size adjectives will be considered in greater detail since this was the only cluster where the statistical analyses resulted in a significant difference between the pairs. In the other cases, some pairs display considerable differences but none of them, possibly because of the smaller sample size, can be considered statistically significant. Despite this lack of statistical significance, a comparison of meanRT between the pairs reveals patterns which not only have theoretical implications for antonymic strength but also for the design of priming paradigms.

##### RELATED-UNRELATED PAIRS

The four related-unrelated pairs (cf. Table 4.5) are all of similar length and should thus allow for reliable comparison of reaction times.

**Table 4.5** *Related-unrelated antonym pairs and their meanRT*

| BASE | ANT 1 | MEAN RT | GOE | ANT 2 | MEAN RT | GOE |
|---|---|---|---|---|---|---|
| happy | sad | 590.00 | 1.215 | unhappy | 546.30 | 1.24 |
| healthy | sick | 659.16 | 1.315 | unhealthy | 530.11 | 1.325 |
| even | odd | 597.77 | 1.3 | uneven | 560.20 | 1.455 |
| married | single | 577.62 | 2.705 | unmarried | 576.00 | 1.965 |

At first glance, it seems that the morphologically related forms are primed more strongly than their unrelated counterparts despite the fact that in three out of the four cases listed above, the unrelated pairs score marginally better in the GOE-rating. The case of *married:unmarried* and *married:single*, which is the case which shows the biggest discrepancy in the GOE-rating in favour of the related pair (1.965 vs. 2.705), will not be discussed further here since the reaction times as well as the standard deviation of the two pairs are identical. This shows that, as far as the data presented here can be considered representative, *married* is as strongly antonymically related to *single* as it is to *unmarried*.

In the other three cases, there is a difference of between 37 and 130ms difference between the related and unrelated pairs. Since repetition priming (cf. among others Harley 2008: 176) is the strongest form of priming and according to Harley (2008: 177) 'recent results suggest that morphological priming is obtained because of morphological structure rather than semantic overlap', it is unsurprising that targets which repeat a large proportion of the prime with only a modifying prefix which is expected in an antonym task prime more strongly in the related than in the unrelated case. The difference in reaction time between the two contrastive pairs in each of the three cases may not be a reflection on the degree of antonymic strength displayed by either combination but instead simply a result of a particularly effective form of priming. The differences in GOE-rating in these three cases were also minimal (cf. Table 4.5 above) and never exceeded 0.15 points difference.

Thus, there is no evidence in this particular data for the hypothesis presented above as the discrepancies in meanRT could easily be explained by the strength of identical priming. This is, however, not the case in the next group of cases.

#### UNRELATED-UNRELATED PAIRS

As in the group above, the word lengths are relatively similar with an average length of four letters, thus there is no expectation of an effect due to word length. Unlike the pairs discussed in the previous section, the pairs in Table 4.6 below are not morphologically related and thus the influence of repetition priming, which may have played a role in the results presented above, is irrelevant here.

**Table 4.6** *Unrelated-unrelated antonym pairs and their meanRT*

| BASE | ANT 1 | MEAN RT | GOE | ANT 2 | MEAN RT | GOE |
|---|---|---|---|---|---|---|
| short | tall | 553.88 | 1.18 | long | 581.10 | 1.215 |
| good | bad | 593.40 | 1.175 | evil | 602.50 | 1.39 |
| hard | easy | 652.00 | 1.55 | soft | 488.77 | 1.605 |
| right | left | 508.44 | 1.305 | wrong | 518.71 | 1.09 |
| sweet | bitter | 664.25 | 2.63 | sour | 587.37 | 2.68 |

The case of the triad *sweet:sour:bitter* generated too many errors (either wrong responses or late responses which were not recorded) and the meanRTs are not representative since not enough subjects answered correctly or within the time frame. The behavioural data thus cannot be used to shed light on the question which of the two pairs is more closely related.

There are two cases in which the results of the experimental task are surprising given the GOE-rating results. In the case of *good:bad* and *good:evil* (cf.

3.4.1.2), one could assume that the former was the pair with the greater degree of associative and antonymic strength. When word association measures, for example the EAT, are consulted, they show that *bad* is the overwhelming response to the stimulus *good* (78/100 participants) wheras *evil* is only mentioned twice. However, since participants are only permitted one response in these association experiments, it cannot be verified how many informants would have listed *evil* as their second choice. Therefore, word association figures like the ones above, while undoubtedly a strong indicator, need to be treated with some caution. Another caveat needs to be issued on the matter of associative measures: associative strength is strongly directional. One lexeme which elicits another very frequently is not necessarily that lexeme's strongest associate since the very strong association may only be evident in one direction (cf. Gross et al. 1989: 100, Hutchison 2003). Nevertheless, most native speakers would, if presented with the choice, presumably select *good:bad* as the more strongly antonymic pair and the GOE-rating results point in the same direction. The meanRT of the two pairs, however, does not show any significant difference.

Another case in which the pairs are surprisingly similar in terms of meanRT is that of the base adjective *right*. *Right:wrong* and *right:left* barely differ in terms of reaction time but do show considerable difference in GOE-rating scores. In this case, associative strength as determined by word association experiments is, if the EAT is considered a reliable source, much more evenly distributed than in the case of *good:bad*/*good:evil*. *Wrong* is the first choice as a response to *right* for 42/100 informants but *left*, with 33/100 responses, is not very far behind. This is more in keeping with the behavioural results than with the GOE-rating where the difference between the two pairs is larger.

The other two sets of pairs display a certain amount of difference between the meanRTs; in the first case, *short:tall* - *short:long*, it is a rather small difference (30ms), which correlates with the amount of difference in the GOE-rating scores which suggest that *short:tall* is the pair with a slightly stronger relation. The EAT results, however, show greater associative strength between *short:long* (42/100) than *short:tall* (14/100). In the second set, *hard:soft* – *hard:easy*, the difference is substantial (163ms) while the GOE-rating shows barely any difference. Associative strength, as predicted by the EAT results, is much greater for *hard:soft* (44/100) while *hard:easy* is only mentioned as a first reaction by one participant. Interestingly, *easy* elicits *hard* as the first reaction 32/100 times while *difficult* ranks second with 20/100. *Hard:easy* is thus one of the pairs where there is a large discrepancy between the strength of association in the different directions. While this ties in with speaker intuitions, it does not match the results of the judgement task.

In the group of the unrelated-unrelated pairs, there are again no statistically significant differences between the amount the base member of the pairs primes its two associates. However, some of the differences are indicative of an initial difference which is not simply a result of greater associative strength.

#### THE SIZE ADJECTIVES

The four size adjectives have been discussed from a number of perspectives in previous chapters. However, there has not yet been a direct comparison between the two pairs which share a base adjective. The experiment included eight pairs since all four adjectives appeared with both possible partners. Direct comparisons will be made between the pairs which are generally considered the 'correct' combinations (cf. 3.4.1.3) as well as between the pairs which share the same base adjective.

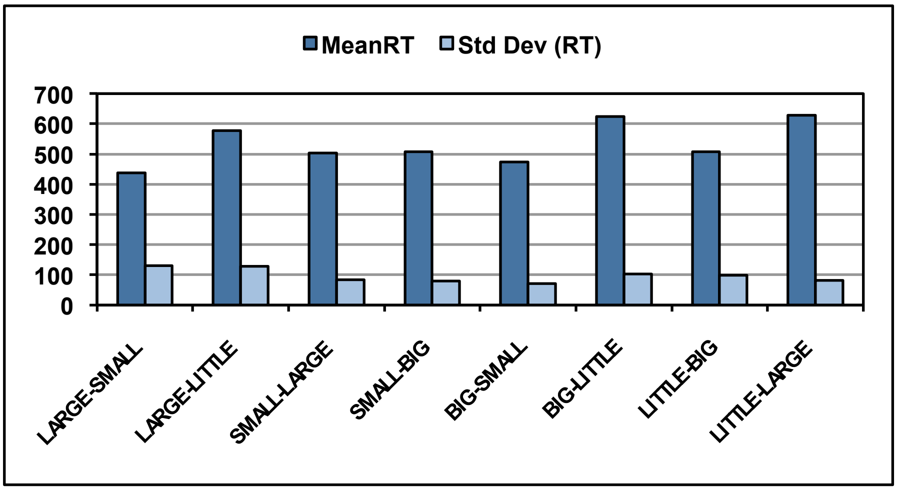



**Figure 4.21** *MeanRT (in ms) for SIZE adjectives in Experiment 2*

The group of size adjectives is the only cluster in this experiment which shows statistically significant differences between pairs which share the same base adjective in a comparison of meanRT. These differences are shown in Figure 4.21 above alongside the mean standard deviation for each item and both measures will be discussed in detail below.

Although some of the differences between items in the previous chapters were considerable, they did not reach a level of statistical significance. This may be, as suggested above, due to the limited sample size used in this study. In case of the size adjectives, fewer errors were made than with the other pairs in this experiment for all mutations of the four adjectives. The standard deviations, as can be seen in Figure 4.21, were also comparatively small for most of the pairs. Table 4.7 shows a comparison of meanRT and GOE-rating for all pairs in this group.

**Table 4.7** *Comparison of MeanRT and GOE-rating scores for SIZE adjectives*

| BASE | ANT 1 | MEAN RT | GOE | ANT 2 | MEAN RT | GOE |
|---|---|---|---|---|---|---|
| little | big | 507.75 | 1.18 | large | 628.44 | 1.95 |
| small | big | 507.88 | 1.35 | large | 504.55 | 1.25 |
| big | small | 474.11 | 1.1 | little | 624.22 | 1.3 |
| large | small | 438.3 | 1.18 | little | 578.14 | 2.15 |

The first observation which can be made here is that, even at first glance, the behavioural results are relatively close to those obtained from the judgement task. There are three pairs which are very close in the GOE-rating and one which is judged to have considerably less antonymic strength (*large:little*). This is reflected in the meanRT measures in the table above where, apart from one important exception (*big:little*) which will be discussed in detail below, *large:little* and *little:large* have by far the longest reaction times. MeanRT is shorter than in Experiment 1 for all pairs in this group, which shows the effectiveness in terms of activation of the priming design compared to simultaneous presentation of both lexemes as a complete word pair. The differences within the group shown in Table 4.7 above are similar to those which can be seen in the results of Experiment 1 (cf. 4.3.3).

The initial step in the analysis of antonymic strength of the size adjectives will be a direct comparison between each of the two pairs which share a first member to investigate whether certain lexemes have a preferred antonymic 'partner' or whether the first member of an antonym pair elicits both possible partners to the same degree. The only lexeme which does not seem to show a strong preference for one of its two possible partners is *small*. It acts as a prime for both *large* and *big* in equal measure. This can be concluded despite the fact that *small:large* is confirmed as an antonym pair slightly faster than *small:big* since the difference in the reaction times between the two combinations is negligible (at 13ms). In every other case (cf. Table 4.7) one combination is recognised as a pair of opposites considerably faster than the other. In the case of *little:big* and *little:large* the difference does not quite reach statistical significance but the former is confirmed much more quickly than the latter.

The last two cases both display statistically significant differences between the two combinations in a t-test ($F(43, 315498) = 12.1618$). In the case of *large:small* and *large:little* ($p \leq 0.0019$), this was expected since the GOE-rating results as well as speaker intuition suggest that *large:little* is not as good an opposition as *large:small*. Thus it seems only logical that *large* should activate *small* faster than *little* and prime it more strongly. This prediction seems to be confirmed by the data. The second case, however, is somewhat surprising and will be discussed in greater detail below since the difference ($p \leq 0.0011$) in GOE-rating of *big:little* and *big:small* does not match the difference in reaction time between these two pairs at all.

The 'correct' opposite partner for *big* is usually said to be *little* (cf. 3.4.1.3) when native speakers of English are asked for their opinion as to which of the size adjectives form the best antonym pairs (cf. also K. J. Miller 1998: 52). Nevertheless, data from the EAT suggests that *small* has, in fact, a better associative connection with *big* than *little* does (scores are 29/100 and 18/100 respectively). As previously shown, a particular distribution of associative strength does not necessarily predict an identical distribution in antonymic strength. In this case, however, the behavioural data supports the data taken from the associative thesaurus. There is a statistically significant difference between *big:small* and *big:little* and the latter pair is the pair with the slowest overall meanRT among the size adjectives despite receiving very similar GOE-rating scores to the other 'good' combinations in the judgement task. Here, the prime seems to generate an inhibitory effect (Harley 2008: 177) instead of facilitating a faster reaction to the target. It seems that *big* primes *small* very strongly and that the strong associative and antonymic bond between those two lexemes inhibits the activation of *little* as a possible opposite partner. Bearing in mind that the subjects are aware they are asked to make decisions on the antonymicity of the pairs they are presented with, the results seem to suggest that *small* is activated as a result of the prime *big* and when the pair shown differs from the expectations of the participant (i.e. *big:little*), the reaction time is slowed down considerably because an unexpected item needs to be selected .

Since the substantial difference between *big:small* and *big:little* is not the only noteworthy discrepancy, the results of the behavioural experiments lead to a re-examination of the previously assumed relationships between the size adjectives. Figure 4.22 below is a new version of Figure 3.6 (3.4.1.3) which includes the insights gained from the behavioural data. This data shows a large effect of directionality in the group of size adjectives. The fact that antonymic relationships are not always symmetrical has been discussed before (cf. 4.4.2.2 – unrelated-unrelated pairs) and the reaction times in this study reflect this asymmetry. While one member of a pair may elicit the other as its primary antonym partner, this is not necessarily the case in the opposite direction; for example, *evil* elicits *good* as its primary opposite whereas *good* elicits *bad* (cf. 3.4.1.3 and 5.2.3.4). The distribution among the size adjectives is slightly more complex since it is not a simple case of one lexeme having two opposites but of connections of varying antonymic strength between four lexemes which are central to a category. A case could also be made for including symmetry of antonymic strength in the list of criteria for good opposition since the argument seems convincing that the higher the reciprocity of elicitation, the more antonymic strength will be attributed to any given pair.

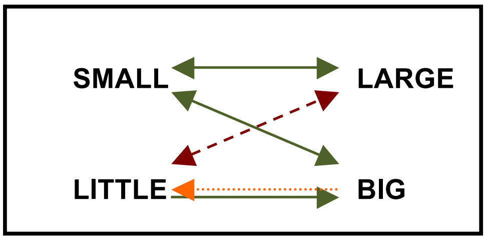



**Figure 4.22** *Relationships between* big, large, small *and* little *revisited*

Green arrows indicate that the antonymic connection between the two lexemes is a primary one whereas orange arrows show secondary opposite partners. The red arrow between *large* and *little* shows the reciprocal poor connection between the two lexemes compared to the other connections in this group. In comparison with many other antonym pairs in this study, however, these two adjectives still fall into the category of good opposition in the judgement task as well as the behavioural experiments.

As can be seen from Figure 4.22, the only reciprocal primary antonymic connection is between *small* and *large*. The primary antonym for *big* is very clearly *small* while *small* elicits *large* before *big* (in behavioural data as well as in the EAT (26/100 *large*; 20/100 *big*)). While *big* leads to *small*, *little* leads to *big* in the first instance and then to *large*. This distribution of primary and secondary links creates an imbalance in a group of opposites which is more often than not said to have two primary opposite pairs (*big:little* and *large:small*) and makes the analysis of lexical opposition here a much more complex matter. In the cases of all primary connections, priming is stronger presumably as a result of both the associative strength and the antonymic strength of the pair being greater than for secondary associations.

The difference in meanRT between *small:large* and *small:big* is insignificant and it can be said that *small* genuinely seems to have two equally strong primary associations, which is a reasonably rare phenomenon among opposites. The central adjectives on the SIZE continuum are a special case among the gradable adjectival antonyms, since there is not one single central antonym pair around which the category is structured but several combinations of the central adjectives which are of approximately equal strength despite different degrees of reciprocity in some of the pairs.

### 4.4.3 Conclusions

The hypothesis underlying Experiment 2 has been confirmed to a certain degree. Overall, the mean reaction times in this task were considerably faster than those in Experiment 1 for most target pairs. This shows that the relatively short display time of 300ms of the first item and a SOA of 450ms do lead to significant priming. Whether this priming is associative or semantic cannot be determined since all target pairs in this experiment have very high associative strength as measured by word association and co-occurrence analyses as well as a large overlap in features and thus a strong semantic (in this case antonymic) relationship (cf. Hutchison 2003). For the hypothesis introduced in 4.4, however, it is not primarily important to distinguish between the relationships which facilitate priming but to investigate the difference the priming paradigm makes firstly to the simultaneous presentation and to individual antonym pairs.

The success of the methodology in some of the cases above, most notably the group of size adjectives, shows that it is possible to determine whether a lexeme has a preferred ‘primary’ partner which is activated automatically by priming the antonym pair with the lexeme in question. This method highlights asymmetries in the reciprocity of antonymic stength in those cases where a lexeme has two possible antonyms. This seems to be related to the issue of sequence in antonym pairs since there are clearly cases where antonymic strength is greater in one direction than in the other (cf. 4.4.2.2), even in pairs which only have one ‘primary’ opposite partner. However, as has been shown above, some lexemes have two equally strongly related ‘primary’ opposites (cf. also Gross et al. 1989: 100).

The analysis of the related-unrelated group showed a tendency for morphologically related lexeme pairs to be recognised faster than those which are not related; this can be put down to the extraordinarily strong priming effect caused by repetition priming (Harley 2008: 176) and it can therefore not be said with any certainty whether the antonymic strength of these items is indeed stronger than those in related-unrelated pairs or whether the RT data is simply an artifact generated by the methodology used. The latter case seems more likely in the context of the analysis of Experiment 1 and the GOE-rating results for the pairs in question.

Overall, it can be said that while the results do not allow for a distinction between the influence of antonymic and associative strength in antonym priming, the data sheds some light on several key issues concerning both appropriate methodology for antonym research, such as the confunding results caused by repetition priming, and the relationship between lexemes which can form more than one antonym pair between them. The results have been particularly strong where the group of size adjectives is concerned and have allowed for an insightful analysis of the complex relationship between those lexemes. In the other groups which were analysed, it was discovered that a larger data set would be needed to turn the observed tendencies into statistically significant results.

## 4.5 Further implications

The intention of the separate analyses carried out in the first half of this chapter was to investigate which individual factors have an influence on antonymic strength and whether there are any differences between the opposite types proposed by Lyons (1977) and Cruse (1986). The analyses above have shown that there are several factors for which the different conditions result in a significant difference in the behavioural results. As expected, GOE-rating and t-score both have very strong effects and these two factors also interact to a certain degree but they do not correlate completely, which was predicted by the results of the judgement task and their comparison to the t-scores in Chapter 3. There are several pairs among the complementaries and converses which obtain results in this study that are at odds with the results of the judgement task (e.g. *king:queen*, *mother:father*). These are items with extremely high associative strength which suggests greater influence of this factor in this type of method. The high degree of associative strength seems to facilitate a faster reaction to the pair than would have been warranted by the antonymic strength displayed in the judgement task.

There are several factors which affect only a subset of the targets in this study. Some of these factors, symmetry for instance, were found to be highly significant within the subset (in this case gradable antonyms). Others, such as gradability and morphological relatedness, did not show much of an effect in the behavioural data. Sequence, the order in which the members of a pair are presented, did not have a significant effect in the whole dataset; however, as can be seen from the discussion in 4.3.1.3, there are certain pairs which are very strongly affected by sequence and which follow the criteria established for sequence in the literature (cf. for example Jones 2002).

The question of interaction between the individual factors was addressed in 4.3.2 (and also 4.3.1) and showed interesting results. There are some factors – those, for example, which apply to only a subset of the opposite pairs – which would not be expected to interact. However, some of the more central criteria were investigated more closely after the inital single-factor analyses and were found to interact very strongly. There is significant interaction between t-score and GOE-rating, GOE-rating and word class and t-score and word class. It shows the large degree of relatedness of these factors and illustrates why antonymy is not a relation determined by a single factor. This ties into the question of whether certain factors allow us to predict whether an antonym pair will be on a particular part of the continuum of antonym canonicity, if such a continuum exists. The evidence for and against a continuum approach to antonym canonicity will be investigated in detail in 6.3 where the results above will be compared with those of Paradis et al. (2009) who used a multi-method approach similar to that employed in the present research. As far as the influence of individual features is concerned, there is a distinct difference between the judgement task and the behavioural task in the influence associative strength seems to have on the results. It seems to be a much more important factor in the lexical decision task and there is a greater correlation between high t-scores and fast reaction times than the GOE-rating results would have predicted. This leads to the conclusion that both associative and antonymic strength play a part in antonym processing and that the weighting of these criteria depends strongly on the type of task which is used to test them. The question which remains to be answered here from a methodological perspective is whether the increased importance of associative strength is simply a by-product of the method or whether it is a genuine factor in antonym judgements. This will be considered in 6.1.1.5 when the relation of antonymy is examined more closely from a theoretical perspective and it will further be addressed from a methodological standpoint in 6.2.2 and 7.2. GOE-rating seems to have the most consistent effect overall on meanRT; however, t-score is also a crucial factor. All other factors contribute to antonymic strength but there is no other single factor which determines as strongly where on the canonicity scale an opposite will be placed. GOE-rating, of course, already includes many of the other factors which determine antonymic strength since those would have played a role in the judgement task which resulted in the GOE-rating scores. It is not possible from the analyses carried out above to distinguish precisely between the influence of antonymic and associative strength.

Further investigation is necessary to address the effects generated by very high t-scores in the behavioural task.

The last question posed at the beginning of the behavioural analysis referred to the influence of antonym type. The analyses have shown that the categories of antonyms and complementaries show very little difference in the overall analysis whereas converses obtain much lower reaction times. Whether this is an indication that converses should not be considered part of the category of opposites (cf. Cruse 1986), or whether it is a reflection of the fact that all converses in the study were nominal or verbal rather than adjectival can not be said with certainty at this stage. Both the word-class aspect and thus the complexity hypothesis introduced at the end of the previous chapter are assumed to play a role in this distribution as well as the fact that the relationship between the two members of a converse pair is not simply antonymous in the canonical sense but is of a more complex nature.

The data gathered and introduced in this and the previous chapter will now be compared to the results of the study on German to further add to the pool of data and determine which of the effects observed in the English data are replicated in German and which, if any, seem to be either specific to English or may have resulted from idiosyncrasies in the data. In addition, the data presented here will be considered further from a theoretical perspective in Chapter 6, which will bring together the different strands of evidence to provide a coherent overview and to further develop some of the theoretical issues raised in Chapter 2.

# 5. A contrastive perspective on opposition: English-German

The opposite of *duck* is *drake*.
Remember that, for heaven's sake!
One's female, and the other's male.
In writing to a *drake*, don't fail
To start your letter off, "Dear Sir."
"Dear Madam" is what *ducks* prefer.
In snowball fights, the opposite
Of *duck*, of course, is *getting hit*.
(Wilbur 2004: 518)

Following the analysis and discussion of the English questionnaire and behavioural data in Chapters 3 and 4, this chapter considers the data collected in the German counterparts to these experiments and compares it with the English results. The comparison of GOE-ratings and behavioural results for German and English opposite pairs of all degrees of antonymicity has several benefits. Firstly, the fact that the same analyses are carried out on a second set of stimuli lends support to the choice of the experimental methods and design as well as providing the basis for a cross-linguistic analysis. One of the key points of this analysis will be the discussion of the question whether certain opposed concepts generate similar GOE-rating judgements in the two languages, which would further corroborate the hypothesis that the lexical relation between the members of a pair is a result of their conceptual relation rather than a cause. Another focus of the analysis is, as in the previous chapters, a closer look at those groups or individual pairs where a large discrepancy can be seen between the German and English pairs and an investigation of possible reasons for these differences. By comparing the antonymic strength of pairs in two languages, it can be established whether all criteria which have been proposed for good antonymy in English also hold in German and whether there is a difference in the extent to which single criteria can influence antonym judgements.

German was chosen as the second language for two reasons, one linguistic and one cultural. Firstly, since both languages are West Germanic in origin, there are enough similarities and shared lexical features between the languages to make the translation of the questionnaire and the behavioural stimuli relatively unproblematic and the encoding of many concepts sufficiently similar. There are, however, also enough differences to make a contrastive analysis both interesting and relevant. For the type of tasks used in this study it was, in the first instance, also important to use two languages spoken in somewhat similar cultures since this allows parallels to be drawn in order to provide satsifying evidence for the hypotheses presented in previous chapters (e.g. the complexity hypothesis in 3.4.3).

The goodness-of-exemplar questionnaire results will be discussed first since they, as in the case of the English study, provide a baseline for the analysis of the behavioural data. A brief general overview of the questionnaire data preceeds the more detailed one-to-one comparison of those word fields (TEMPERATURE, MERIT, SIZE and GENDER) which were already considered of particular interest in the previous analyses. A return to these by now well-known examples shows illuminating similarities and discrepancies between the German and English data. The last section of the questionnaire data highlights two particularly interesting cases before the behavioural data is considered and added to the analysis to determine whether the same factors account for antonymic strength in English and German.

## 5.1 A brief overview of the German data[1]

Before the data is analysed in detail, an overview of the general results of the goodness-of-exemplar rating questionnaire is given. The four different German GOE-rating questionnaires were each completed by 40 participants who were recruited via e-mail. The subjects were between the ages of 19 and 60 with the majority in their mid-twenties. All participants were native speakers of German and most (137 out of 160) were Bavarian but, as in the English questionnaire, dialectal differences were not taken into account.

The methodology used to collect the German data is identical to that used for the English questionnaire. The collection software is the same, as are the number and structure of word-pairs and the number of participants for each of the four questionnaires. Therefore, no methodological details will be presented here (for an explanation of the methods used cf. 3.2). The choice of the German word pairs, however, needs further explanation.

The German questionnaire was constructed to be as similar as possible to its English counterpart and most items are direct translations of the English word pairs (e.g. *hot:cold – heiß:kalt*; *mother:father – Mutter:Vater)*. 38 English pairs, however, had to be replaced with non-matching German pairs since not all of the pairs

[1] Translations for German pairs will be given in the text for individual examples. Translations of larger sets of German data (e.g. attribute listing data) are provided in Appendix 11.

contained in the English questionnaire have a German translation equivalent.[2] For example, on the size continuum English has, as discussed in 3.4.1.3, four central lexemes which can be paired in a variety of ways (*big, little, large, small*) whereas German only has two lexemes (*groß* and *klein*) which not only encode the meanings of the four terms above but also part of the meanings of English *short* and *tall*. However, German does have a similar occurrence of a quartet of lexemes which can be paired in several different ways: *schwer*, *einfach*, *schwierig* and *leicht* ('heavy/difficult', 'easy', 'difficult', 'light/easy'). Thus, redundant combinations of the English quartet were replaced by similar German ones. The relationships between the lexemes in the German quartet are not exactly the same as those in the English one; this will be discussed in more detail in 5.2.2.

Other replacements include that of the triplet *organised*, *unorganised* and *disorganised*; the German questionnaire has the morphologically similar triplet *moralisch*, *unmoralisch* and *amoralisch* ('moral', both 'immoral') in its place. This policy of substituting structurally similar pairs or clusters of pairs for those English pairs where a semantic equivalent cannot be found was adopted throughout. Thus the German questionnaire contains almost exactly the same number of all types of opposite pairs as the English questionnaire. Each of the main groups and their overall results will be presented in more detail below.

The co-occurrence data for the German pairs was extracted from the corpus of the Institut für Deutsche Sprache (Mannheim) using their web-based user-interface COSMAS II.[3] This corpus is the biggest freely available German-language corpus and the part used for the present research[4] is over two billion word-forms (2,291,515,012 tokens) in size. The same t-test tool used for the English data was also used for the German data but the t-scores obtained are comparatively larger, due to the larger number of overall occurrences as a result of the significantly bigger corpus. This is a methodological difficulty which has to be taken into account since the number of relatively simple statistical tests available for the range of data obtained through the corpus studies is limited and there will be a certain amount of distortion whichever test is used. Thus, it needs to be borne in mind when considering the t-score data below that the somewhat inflated t-scores for the German pairs are not always indicative of more frequent co-occurrence of the two lexemes. However, since the main comparison which is made is that of the GOE-

---

[2] For a list of all German pairs and their English equivalents, see Appendix 2. All items were translated by the experimenter and the completed list was checked by two German linguists to ensure the acceptability of the word pairs.
[3] www.ids-mannheim.de and https://cosmas2.ids-mannheim.de/cosmas2-web/
[4] Public part of the Archiv der Geschriebenen Sprache (W – öffentlich) http://www.ids-mannheim.de/cosmas2/projekt/referenz/archive.html (last accessed 03.05.2010)

ratings and the comparison of t-score and GOE-rating is mainly an intra-linguistic measure, this does not affect the validity of the data presented below.

### 5.1.1 Morphological relatedness

The German questionnaire contains 28 morphologically related pairs as well as 35 non-antonym pairs (e.g. *traurig:unglücklich*; 'sad:unhappy'). All 210 pairs are subdivided into four groups by their goodness-of-exemplar rating score (or antonymic strength judgement): group I (1.0 - 1.79) contains excellent antonyms and is the largest with 61 antonym pairs, group II (1.80 – 2.99) contains 48 pairs, group III (3.0-4.99) 58 pairs and group IV (4.0 – 7.0) 45. This division is remarkably similar to that found in the English questionnaire for which the group sizes were as follows: group I 64 items, group II 41 items, group III 57 items and group IV 48 items. The fact that the overall results, regardless of the ratings given for individual pairs, are this consistent indicates that antonymic judgements are similar to a certain extent across languages regardless of lexical encoding.[5]

Similar to the English data, the highest scoring pairs are morphologically unrelated (e.g. *neu:alt, oben:unten, groß:klein;* 'new:old', 'up:down', 'big:small') but almost all morphologically related pairs fall into groups I and II (with the exception of *packen:auspacken* and *lernen:verlernen;* 'pack:unpack', 'learn:unlearn') which are, for different reasons (unnatural pairing in the former example), not ideally chosen. The vast majority of the pairs in group I are adjectival opposites with only four (out of a total of 49) nominal (*Zwerg:Riese, Tod:Leben, Mann:Frau* and *Berg:Tal;* 'dwarf:giant, 'death:life', 'man:woman', 'mountain:valley') and five (out of 24) verbal pairs (*nehmen:geben, ausziehen:anziehen, erlauben:verbieten, kaufen:verkaufen* and *fragen:antworten*; 'take:give', 'dress:undress', 'allow:forbid', 'buy:sell', 'ask:answer'). There are roughly the same number of non-adjectival pairs in group I of the English GOE-rating but the vast majority of those pairs are verbal (two nominal and eight verbal pairs). Apart from the smaller numbers of pairs in group I, verbal and nominal pairs are fairly equally represented throughout the other groups (nouns: 13/18/14; verbs: 7/6/6).

[5] For a table of all German results (GOE-raing and t-scores), see Appendix 9.

### 5.1.2 Types of opposition

The questionnaire contained 209 pairs in total, of which 42 were designed to be non-antonymic. Some of these non-antonymic pairs were synonyms (e.g. *clever:helle*; 'clever:bright'), meronyms (e.g. *Buch:Seite*; 'book:page') or co-hyponyms (e.g. *blau:orange*; 'blue:orange') while others were associatively related (e.g. *Krone:König*; 'king:crown'). All pairs presented in the questionnaire displayed some degree of semantic or associative relatedness since the decision the participants were required to make was a meta-linguistic one focusing on a particular lexical relation rather than one requiring subjects to distinguish between related and unrelated word pairs. In analogy to the English questionnaire, type of antonymy (cf. 2.1) was a criterion in the choice of word pairs. Four types of antonymy are represented in the questionnaire: gradable opposites, non-gradable opposites, converses and spatial opposites. Each of the groups will be introduced briefly below before the main analysis in 5.2.

#### GRADABLE OPPOSITES

The group of gradable opposites is the largest, consisting of 66 word pairs. Gradability is an important criterion in the literature (cf. 2.1.2) and a large proportion of the highest scoring pairs (e.g. *hell:dunkel*, *alt:jung*, *groß:klein* and *heiß:kalt;* 'light:dark, 'old:young', 'big:small' and 'hot:cold') are examples of gradable lexical opposition. Many of the non-canonical gradable pairs are constructed around a central, canonical pair to form a word field of related pairs testing the criteria of symmetry and purity of opposition.[6] The German pairs are matched as closely as possible to the English pairs but there are certain lexical fields which are markedly different in terms of the number of lexemes available to describe certain properties. The fields which will be discussed in greater detail below are those which have already been presented in the analysis of the English questionnaire and behavioural data: TEMPERATURE (cf. 3.4.1.1), MERIT (cf. 3.4.1.2) and SIZE (cf. 3.4.1.3).

#### NON-GRADABLE OPPOSITES

Non-gradable opposites can be divided into two sub-categories: **complementaries** (e.g. *male:female*) and **incompatibles** (e.g. *cat:dog*). The questionnaire contains 60 complementary pairs and 15 incompatibles among the opposite pairs. It is not always easy to distinguish between these two categories and gradable opposition

[6] The method used to construct these opposite pairs is illustrated in 3.1.2.1.

since some pairs have an absolute and a gradable member and some, like *clean:dirty*, seem to be borderline gradable (cf. 2.1.2; Cruse 1986: 203). However, since this is mostly a formal (and logical) decision which is not reflected in the data, all pairs containing an absolute member are included in the category of non-gradable opposition for ease of classification. The non-gradable pairs, like the gradable ones, cover the whole range of judgement results since some of them are constructed to be better than others. The gender cluster will be discussed in more detail below and, once again, includes pairs which have been specifically designed to examine the criterion of purity of opposition.

#### CONVERSES

There are eighteen converse pairs in the German questionnaire which are divided into nominal (8) and verbal (10) converses for the purpose of the analysis. Both groups will be discussed briefly here and later compared to results of the English questionnaire (cf. 5.2.2) to determine any siginificant differences in the judgement of antonymic strength.

The verbal converses range from 1.405 (*geben:nehmen;* 'give:take') to 5.63 (*jagen:fliehen;* 'chase:flee') in goodness-of-exemplar ratings. These scores cover almost the whole range of the Likert-scale used in the questionnaire and indicate that converses, like other types of antonymy, form a prototype category which contains better and worse examples of verbal converses.

With the exception of *Mann:Frau* which was included as a translation equivalent for *husband:wife*, all nominal converses cluster between 2.83 (*Opfer:Mörder;* 'victim:murderer') and 3.64 (*Eltern:Kinder;* 'parents:children'). The fact that *Mann:Frau* does not fit the pattern displayed by the other nominal converses is unsurprising since both German lexemes incorporate the meanings of *man* and *husband* and *woman* and *wife* respectively thus making them much less of a converse pair since the complementary nature of *man:woman* appears to be more salient here.

#### DIRECTIONAL OPPOSITES

The category of directional opposites is very small and only contains eight pairs. However, it was included to investigate the question whether spatial opposites, which are generally considered very good examples of opposition, are similarly "good" examples in both languages. As this category will not be discussed separately below, the data is presented in Table 5.1.

**Table 5.1** *Comparison of German and English spatial opposites (by GOE-rating)*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| up | down | 1.04 | 50.301 | oben | unten | 1.03 | 122.689 |
| top | bottom | 1.115 | 21.390 | rechts | links | 1.205 | 189.493 |
| far | near | 1.19 | 8.437 | hinaus | hinein | 1.415 | 18.577 |
| out | in | 1.24 | -3.076 | hinter | vor | 1.53 | 60.22 |
| right | left | 1.305 | 33.41 | hoch | runter | 1.675 | 29.629 |
| over | under | 1.365 | -2.087 | über | unter | 1.75 | 30.308 |
| behind | in front | 1.675 | 7.032 | weit | nah | 1.87 | 14.363 |
| next to | opposite | 5.19 | 1.652 | neben | gegenüber | 4.68 | 1.385 |

Six out of the eight English pairs have an equivalent German pair, *top:bottom* does not have any corresponding pairs in German while *up:down* has two: *oben:unten* and *hoch:runter*. The former is used only for as a stative spatial indicator whereas the latter is used when the direction of movement is of primary importance. Compare the examples in 5.1 which illustrate this distinction.

(5.1) *Da* ***oben*** *ist ein Ballon!* vs. *Der Ballon steigt* ***hoch****!*
There up be.3.SG.PRS a.ART.M balloon.NOM The.ART.M balloon.NOM rise.3.SG.PRS up
'There's a balloon **up** there!' vs. 'The balloon is going **up**!'

Other antonym pairs which encode the directional aspect of *up:down* in German are the more formal *hinauf:hinunter* or *herauf:herunter* as well as the informal *rauf:runter*. The decision to use the combination *hoch:runter* was a deliberate one since it seemed less informal than *rauf:runter* but more familiar than the more formal variants. The opposite decision was made with *hinaus:hinein* which encodes exactly the same meaning as *hinaus:herein* as well as the more informal counterpart *raus:rein*. This was a merely subjective judgement by the experimenter since the focus of this study is not on the variants of spatial terms but it would be interesting to investigate this phenomenon in more detail.

It is curious that *up:down* and *oben:unten* score highest among the spatial opposites (*oben:unten* is, in fact, the highest scoring pair in the questionnaire whereas its English equivalent is rated second) while the pair which encodes the other senses of *up:down* is ranked considerably lower. This may be due to the large choice of different lexical items which can be used to encode these concepts. This weakens the antonymic strength of that particular pair. It is, however, by no means a low-scoring pair.

The data above show that the range of judgement ratings is very similar in English and German. Overall, the English pairs score slightly lower (i.e. better) than the German ones with all pairs apart from *next to:opposite* being in group I. Frequency of co-occurrence does not always correspond to the judgement rating. If this were the case, the English pair *far:near* and the German *hinein:hinaus* should both be significantly lower in the ranking of spatial pairs. There is, however, an obvious correlation between frequency and GOE-rating (as also discussed in 4.3.1.1 and 4.3.1.2).

The pairs which exhibit the largest differences in GOE-rating are *far:near – weit:nah* and *over:under – über:unter*. It is difficult to find any reasons for this discrepancy. However, these differences are not particularly large and are not as significant as those which are discussed in the following section.

### 5.1.3 Significant differences in the GOE-rating

The analysis below will focus on differences and similarities in GOE ratings and frequency counts of those groups of antonym pairs which were discussed in detail in the preceding chapters. Before considering those examples, some individual cases which show a marked difference in their GOE rating in the two languages will be introduced. In total, there were only 22 pairs which showed a difference greater than 1.0 between the two questionnaires. Of those, 14 fall into one of the groups which will be considered in detail. Four nominal pairs fall into the category of deliberate borderline antonymy (*town:country – Stadt:Land*, *fish:fowl – Fisch:Fleisch*, *dog:cat – Hund:Katze* and *coffee:tea – Kaffee:Tee*) while the remaining four do not seem to share any common characteristics (*ask:answer – fragen:antworten*, *bland:hot – fad:scharf*, *present:future – Gegenwart:Vergangenheit, imprecise:exact – unpräzise: exakt*).

Interestingly, in the case of the nominal borderline examples above, German speakers always considered those pairs better examples of antonymy than their English counterparts considered the corresponding English pairs whereas German subjects rated asymmetrical non-canonical pairings of gradable adjectives (e.g. *imprecise:exact – unpräzise:exakt* as well as some which will be considered in 5.2.3) as less antonymical than English subjects. Possible reasons for these general tendencies are difficult to determine while individual cases may well be relatively easy to explain. One example here is that of *dog:cat*. German has an idiomatic expression '*wie Hund und Katz*' which denotes two people who do not get along at all and fight incessantly. It is possible that this relatively common expression brings the two lexemes and their attached concepts into a closer antonymic relationship than they would otherwise have. English does have the same expression but it may be less frequent. This nevertheless makes the above hypothesis less convincing.

In *bland:hot* vs. *fad:scharf*, the fact that *fad* is a dialectal lexeme (southern German) may have had an impact on the overall rating since while the majority of participants were from southern Germany, there were also a considerable number from the northern parts of the country. *Fad* is synonymous with *boring* and is therefore much more generally used than *bland* in English as it can refer to people, activities and also, of course, taste. Both the dialectal and semantic range factors may have contributed to its lower ranking in the judgement task. There was, however, no other lexeme which was considered an appropriate opposite to *scharf* (other than *mild* 'mild' which was already in the study in another sub-sense) which is why the non-standard lexical item was chosen.

## 5.2 Contrastive analysis

After this brief overview which outlines the similarities between the English and German data, a selection of cases will be discussed individually. The first section considers the factor of morphological relatedness in more detail; firstly in isolation and then, in the second part of 5.2.1, in the context of nominal and verbal converses since a relatively high proportion of converses are morphologically related in one or the other language. The second part of the contrastive analysis focuses on the comparison of certain lexical fields in German and English. After the quartets briefly mentioned in the introduction of this chapter have been discussed, the word fields of TEMPERATURE, SIZE and MERIT are, once again, treated in depth to illustrate the differences and similarities in the antonymic relations in each field.

### 5.2.1 Differences in morphological relatedness

The questionnaires contained, by design, a number of pairs which were morphologically related in one language and unrelated in the other as well as the triplet pairs discussed in 4.4.2 which result in one morphologically related pair and one unrelated one (*happy:unhappy* and *happy:sad*). The analysis below will take all of these pairs into account to determine whether there are any differences in the ratings depending on morphological relatedness alone as this is the only difference between those lexemes.

Among the converses in Table 5.2, the two nominal converse pairs both differ in terms of their GOE-rating and in both the better score is achieved by the morphologically related pair. However, in the second case (*employer:employee)*, the difference is much smaller than in the first (*landlord:tenant*). The verbal converse pairs, *rent:let*, *buy:sell* and *borrow:lend* and their German equivalents, present a much less clear-cut case. In the only widely accepted converse example, *buy:sell*, the morphologically related German pair scores worse than the unrelated English equivalent. These pairs will be discussed more thoroughly in the following section (5.2.2).

**Table 5.2** *Cross-linguistic differences in morphological relatedness (by type)*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| landlord | tenant | 3.465 | 18.37 | Vermieter | Mieter | 2.515 | 48.741 |
| employer | employee | 2.85 | 10.39 | Chef | Angestellter | 3.075 | 8.206 |
| let | rent | 5.615 | 3.26 | mieten | vermieten | 2.45 | 6.193 |
| sell | buy | 1.67 | 13.99 | verkaufen | kaufen | 1.74 | 44.152 |
| borrow | lend | 2.45 | 4.67 | verleihen | ausleihen | 4.15 | 1.833 |
| king | queen | 4.775 | 15.23 | König | Königin | 3.655 | 46.303 |
| happy | sad | 1.215 | 5.417 | glücklich | unglücklich | 1.29 | 13.517 |
| interested | uninterested | 1.415 | 0.939 | interessiert | uninteressiert | 1.565 | 0.92 |
| interested | disinterested | 2.075 | 1.3608 | interessiert | desinteressiert | 1.52 | 1.621 |
| sick | healthy | 1.315 | 2.745 | krank | gesund | 1.22 | 25.387 |
| unhealthy | healthy | 1.325 | 1.975 | ungesund | gesund | 1.54 | 7.972 |
| unhappy | happy | 1.24 | 3.869 | unglücklich | glücklich | 1.275 | 13.517 |

*King:queen* and *König:Königin* score very differently in the antonymy judgement task and the related pair is judged to be a significantly better case of antonymy than the unrelated one (4.775 vs. 3.655). Since there is only one example in the data this can, of course, not be used to draw any conclusions.

Among the other pairs in the table, there seems to be no considerable difference between the related and unrelated pairs which could indicate a pattern. In general, as already mentioned in 5.1.1, the top group of pairs in both languages is made up of morphologically unrelated items whereas no adjectival morphologically related pair (apart from pseudo-antonyms like *easy:uneasy* or *hilflos:hilfreich*) scores higher than 2.075 (*interested:disinterested*).

### 5.2.2 Converses contrasted

Most of the German converse pairs are very similar in GOE-rating score to their English counterparts and there are very few where the difference is greater than 0.5 on the 7-point scale. As already mentioned in 5.1, the difference in pattern between nominal and verbal converses which was observed in the English data is also in evidence in the German instances of converse pairs. Nominal converses form a cluster which spans roughly one point of the scale whereas the judgements of antonymic strength for verbal converses are much more wide-ranging.

Despite considerable differences in t-score between many of the corresponding verbal converse pairs, the judgement of antonymic strength is remarkably consistent. *Give:take*, for example, has a t-score of 5.356, which is among the lower scores in this group, compared to a t-score of 45.599 for *geben:nehmen* which is the second highest of the German t-scores. Their GOE-ratings are, at 1.49 and 1.405 respectively, very similar. This, once again, shows that while frequency of co-occurrence correlates with high antonymic strength, the influence and direction of that dependence is not at all clear from data such as in table 5.3 below.

As can be seen from Table 5.3, there are two pairs which display a slightly different GOE-rating score: *defend:attack* and *hunt:escape*. The former is judged to possess a relatively high degree of antonymic strength in both languages but, despite a higher t-score, the German *verteidigen:angreifen* scores slightly lower. *Hunt:escape* scores comparatively low in general and is one of the examples which were created to test the boundaries of antonymic strength in converses. It is not a conventional opposite pair but neverless encodes a converse relation. The German counterpart, *jagen:ausbrechen*, also encodes the same relation but *ausbrechen* is slightly narrower in meaning than *escape*. Its usage is more literal – escaping from a confined space – than that of *escape,* which is frequently also used metaphorically, and will usually require an animate subject.

**Table 5.3** *Cross-linguistic comparison of verbal converses*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| sell | buy | 1.67 | 13.99 | verkaufen | kaufen | 1.74 | 44.152 |
| give | take | 1.49 | 5.356 | geben | nehmen | 1.405 | 45.599 |
| defend | attack | 1.44 | 3.07 | verteidigen | angreifen | 1.98 | 8.886 |
| come | go | 1.785 | 15.74 | kommen | gehen | 1.885 | 64.73 |
| borrow | lend | 2.45 | 4.67 | verleihen | ausleihen | 4.15 | 1.833 |
| chase | flee | 3.46 | 0 | jagen | fliehen | 3.44 | 1.812 |
| steal | donate | 4.14 | 0 | stehlen | spenden | 4.665 | -0.373 |
| hunt | escape | 4.765 | 0 | jagen | ausbrechen | 5.63 | 0.838 |
| purchase | trade | 5.305 | 0.985 | erwerben | handeln | 5.59 | 2.005 |
| let | rent | 5.615 | 3.26 | mieten | vermieten | 2.45 | 6.193 |

The two pairs which display the clearest difference are *borrow:lend* and *let:rent*. They are, in fact, opposite manifestations of the same phenomenon. In the case of *borrow:lend* and *verleihen:ausleihen*, English encodes each direction of this converse opposition in a separate lexeme whereas the distinction made by prefixation in German is somewhat artificial since the German lexeme *leihen* (and its prefixed form *ausleihen*) can mean both 'borrow' and 'lend'. In the case of *let:rent* and *mieten:vermieten*, German uses two lexemes while in English, speakers have two lexemes at their disposal but, much like in the case of *leihen* in German, they can be used interchangeably. Both examples will be discussed in turn.

The pair *lend:borrow* is generally considered an opposite pair by speakers of English but not one of the canonical pairs due to its converse nature. There are some dialect areas in the UK and other English-speaking areas (e.g. the USA) which use either lexeme for both concepts (cf. Example 5.2) but in most varieties, including Standard British English, the separation is fairly rigid and 'mistakes' are very noticeable.[7]

(5.2) North West Wales

***Borrow** me a quid* vs. *I **lent** it off him.*

This causes difficulties for native speakers of German who are learning English since German uses one lexeme for both directions of the process: German does have two lexemes to cover this word field, but both can be used to denote *lend* **and** *borrow*, namely ModG *borgen*, which derives from the same root as *borrow* (OHG *borgên*) and *leihen* which can be traced back to OHG *līhan* (MHG *līhen*). *Lend* is not etymologically related to *leihen* but comes from OE *lænan*, which comes from the OE noun *læn.* German has a derivate from the same nominal root via OHG *lêhanôn* (MHG *lêhenen*) which is *lehnen* in ModG meaning 'to enfeoff.'[8] *Lehnen*, as a verb, is very seldom used and is not considered a lexical choice for the concepts under discussion.

(5.3) *lend/borrow* vs. *leihen/borgen*

| | | | | | |
|---|---|---|---|---|---|
| *Kannst* | *du* | *mir* | *das* | *Buch* | ***borgen/leihen**?* |
| Can.2.SG.PRS | you.NOM | me.DAT | the.ART.ACC.N | book.ACC | lend.INF |
| 'Can you lend me the book?' | | | | | |
| *Kann* | *ich* | *mir* | *das* | *Buch* | ***borgen/leihen**?* |
| Can.1.SG.PRS | I.NOM | me.DAT | the.ART.ACC.N | book.ACC | borrow.INF |
| 'Can I borrow the book?' | | | | | |

From the examples in 5.3 it can be seen that both *leihen* and *borgen* can be used for either process. The distinction between the two is dependent on register and

[7] Both the OALD and the LDCE make a particular point of indicating the correct usage of *borrow* and *lend*.

[8] 'To invest with a fief; to put (a person) in possession of the fee-simple or fee-tail of lands, tenements, etc.' (www.oed.com - 22.12.2009)

region; *borgen* is more colloquial and used mainly in northern Germany, whereas *leihen* is the more standard form which is used throughout the German-speaking area. As can be seen from the examples above, there is little room for ambiguity since the syntax of the sentence makes clear which meaning is intended. However, this is also the case in English which nevertheless uses two lexemes. Interestingly, both German lexemes can be modified morphologically by adding the prefix *aus-*, but this does not disambiguate between the two senses since *ausleihen* can still be used for both *lend* and *borrow* (cf. 5.4).

(5.4)

| | | | | |
|---|---|---|---|---|
| *Kann* | *ich* | *mir* | *das* | ***ausleihen****?* |
| Can.1SG.PRS | I.NOM | me.DAT | this.ACC.N | borrow.INF |

'Can I borrow this?'

| | | | | | |
|---|---|---|---|---|---|
| *Ich* | *habe* | *dir* | *das* | *Buch* | ***ausgeliehen****.* |
| I.NOM | have.1SG.PRS | you.DAT | the.ART.N | book.ACC | lend.PST.PTCP |

'I have lent you the book.'

The only alteration which makes it completely clear which direction of the exchange the speaker is focussing on is the addition of *ver-*. This is only an option with *leihen,* but the resulting form *verleihen* is not used very frequently. The corresponding noun, *Verleih*, however, for a rental agency (cars, skis etc.) or for the process of lending, is extremely common.

The pair *rent:let* operates in the same way but is the reverse case. *Rent* could conceivably form an antonym pair with *let,* which covers a similar semantic area, but since neither is completely restricted to either direction of the 'letting/renting exchange' this antonymic pairing is rather weak, a hypothesis which has been substantiated by the low antonymic strength this pair displayed in the GOE-rating questionnaire. *Let* seems more restricted in usage but, as was shown in the case of *ausleihen:verleihen* above, only one of the two lexemes needs to be bi-directional for the opposition to be weakened. In the case of *rent*, the directionality is occasionally made clear by forming a prepositional construction with *out* – *to rent out*. However, as the examples in (5.5) show, it is perfectly possible for *rent* to operate on its own for both perspectives which could have been encoded by a converse antonym pair.

(5.5) *rent*

*He* ***rented*** *the cottage to them.* vs. *They* ***rented*** *the cottage from him.*

Sense 1: *trans.* To let (property) for rent or payment; to hire *out* [sic] to someone. Freq. with the person as indirect object. (c1447/1546)

Sense 2: *trans.* To pay rent for (land, buildings, etc.); to take possession of, hold, occupy, or use, by payment of rent. (1530)

In German, the situation is different; *mieten:vermieten* form a morphologically related antonym pair and the directionality is signalled by the prefix *ver-* which is often, but not exclusively, a marker of converseness (e.g. *kaufen:verkaufen* 'buy:sell'). There is no flexibility in the usage of either *mieten* or *vermieten*; each covers its own part of the relevant semantic field, and using one for the other would have much the same effect as it would to use *buy* and *sell* the 'wrong' way round.

(5.6) *mieten* vs. *vermieten*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Er* | ***vermietete*** | *seine* | *Wohnung*. | vs. | *Sie* | ***mietete*** | *ein* | *Ferienhaus.* |
| He.NOM | rent.3SG.IMP | his.ACC.F | flat.ACC | | She.NOM | rent.3SG.IMP | a.ACC.N | holiday house.ACC |
| 'He rented his flat.' | | | | vs. | 'She rented a holiday home.' | | | |

This pair is readily recognised as an antonym pair by speakers of German (as much as any converse pair in the language) and does not encode weaker opposition than any other pairs of its kind. Thus, it seems plausible that speakers perceive concepts as standing in weaker opposition when they are antagonymically[9] encoded than when the same concepts are encoded antonymically. This is a hypothesis which will be explored in depth in 6.1.1.1.

Among the nominal converses, the most notable difference between German and English has already been mentioned: the GOE-rating of *husband:wife* does not correspond to its German translation equivalent *Mann:Frau*. The explanation for this discrepancy is simple. German does not have words used to exclusively denote the male and female members in a marital relationship. Instead, the words for 'man' and 'woman', *Mann* and *Frau*, are used to refer to 'husband' and 'wife'. The cognate of *wife*, German *Weib* (OE. *wíf*; OHG. *wîp*) is now very old-fashioned and most frequently used with negative connotations while the derived adjective, *weiblich* 'female, feminine', is neutral (or mostly positive).

**Table 5.4** *Cross-linguistic comparison of nominal converses (by English GOE)*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| employer | employee | 2.85 | 10.39 | Chef | Angestellter | 3.075 | 8.206 |
| pupil | teacher | 3.15 | 11.40 | Schüler | Lehrer | 2.905 | 167.193 |
| husband | wife | 3.175 | 28.01 | Mann | Frau | 1.5 | 185.818 |
| victim | murderer | 3.225 | 3.11 | Opfer | Mörder | 2.83 | 24.639 |
| master | apprentice | 3.225 | 2.17 | Meister | Lehrling | 3.045 | 16.06 |
| doctor | patient | 3.375 | 9.76 | Doktor | Patient | 2.96 | 6.824 |
| landlord | tenant | 3.465 | 18.37 | Vermieter | Mieter | 2.515 | 48.741 |
| student | teacher | 3.555 | 9.46 | Student | Dozent | 3.53 | 5.661 |
| parent | child | 3.59 | 15.66 | Eltern | Kinder | 3.645 | 250.239 |

[9] An antagonym, as opposed to an antonym, is a lexeme which has two opposing sub-senses (e.g. *dust* ('to cover in dust' or 'to remove dust')). For work on antagonyms, see Lutzeier 1997 and 2001.

Table 5.4 shows that the only other notable difference is that between *landlord:tenant* and *Vermieter:Mieter*. The German pair is rated as having significantly higher antonymic strength than its English equivalent. Looking back at the discussion of the verbal converse pair *mieten:vermieten*, it seems logical to assume that the existence of a morphologically related verbal converse pair strengthens the oppsition in the corresponding nominal pair. In English, no such equivalent exists and therefore the opposition between *landlord* and *tenant* does not have this additional strengthening factor. The question whether the polysemy of *landlord* (also 'publican') has any influence on antonymic strength is more difficult to determine.

The exceptions which have been discussed aside, both nominal and verbal converses show very strong similarities in antonymic strength between English and German despite differences in frequency of co-occurrence.

### 5.2.3 Contrastive word fields

To examine the extent of the impact of criteria such as symmetry, distance from the mid-point of a scale, and purity of opposition have on the judgement of gradable antonym pairs, the three clusters of opposite pairs used in the English questionnaire were translated into German.[10] There are several cases in which German does not have the same degree of lexical variety as English and thus some of the English pairs do not have a direct German counterpart (cf. 5.1).

As indicated in 5.1, the relationship between the four central size adjectives which was discussed in 3.4.1.3 cannot be transposed directly into German since there are only two adjectives in German (*klein* and *groß*) which cover the same semantic range as the four in English and, in addition to that, parts of the semantic range of *short* and *tall*. However, since the relationships between those four adjectives in English and the pairings they form are very interesting, a group of adjectives which display similar features was substituted in the German questionnaire. This quartet will be discussed in 5.2.2.4, after an analysis of the English-German equivalents in the fields of TEMPERATURE, SIZE and MERIT.

#### *5.2.3.1 The TEMPERATURE scale*

Nine opposite pairs in the German questionnaire are clustered around *heiß:kalt*. All but one of the English pairs on the TEMPERATURE scale have a German equivalent;

[10] All translations were carried out by the researcher and verified by other German native speakers.

*lukewarm:tepid* does not have a corresponding pair since both of these lexemes are encoded in just one, *lauwarm*, in German. Furthermore, one pair, *hot:bland*, does have a German equivalent (*scharf:fad*) but this pair does not include any of the temperature adjectives and will thus not be discussed. It is only included in the table below to give an overview of the results of the complete dataset.

**Table 5.5** *Cross-linguistic comparison of pairs on the* TEMPERATURE *scale*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| hot | cold | 1.13 | 20.32 | heiß | kalt | 1.08 | 30.702 |
| warm | cool | 1.84 | 5.78 | warm | kühl | 2.755 | 13.212 |
| freezing | boiling | 1.975 | 2.22 | eisig | kochend | 2.855 | 0 |
| hot | cool | 2.68 | 5.22 | heiß | kühl | 3.365 | 9.628 |
| chilly | warm | 2.265 | 1.34 | frostig | warm | 3.355 | 2.948 |
| cold | mild | 4.815 | 1.31 | kalt | mild | 4.565 | 6.143 |
| lukewarm | tepid | 6.215 | 0.99 | | | | |
| chilly | steaming | 3.565 | 0 | kühl | brütend | 4.705 | 0.982 |
| cold | friendly | 4.015 | -0.84 | kalt | freundlich | 4.775 | 6.175 |
| bland | hot | 4.855 | 0 | fad | scharf | 3.855 | 1.561 |

Thus, eight pairs remain for discussion. In terms of frequency of co-occurrence, Table 5.5 shows that the t-scores for the English and German pairs are relatively similar. However, there are only three pairs which are close in the GOE-rating results: the base pairs *hot:cold* and *heiß:kalt* which are among the highest-scoring pairs in the questionnaires, the asymmetric pairs *cold:mild* and *kalt:mild* which are unanimously seen as fairly poor examples of lexical opposition, and the pairs *cold:friendly* and *kalt:freundlich* which make use of the polysemous nature of *cold*/*kalt*, cross the word field boundary and are thus not part of the temperature scale.

The five remaining pairs differ in their inter-linguistic GOE-rating to varying degrees and the German pairs are invariably judged to be worse opposites than their English equivalents. The difference in the symmetrical pairs *warm:cool* – *warm:kühl* and *boiling:freezing* – *kochend:eisig* will be considered first before moving on to the discussion of the remaining asymmetrical pairs.

*Warm:kühl* (2.755) is judged to be less strongly antonymically associated than the English pair *warm:cool* (1.84). The latter is almost in the top group of very high antonymic strength whereas the former is at the bottom of the second group (good opposition). In German, the asymmetric pair *warm:kalt,* which was not included in the questionnaire, is more likely to be used instead of *warm:kühl*. *Kühl* seems to be more restricted in range than the English lexeme *cool* and this, in conjunction with there being another pair which covers an almost identical part of the temperature scale, would be more than sufficient to weaken the antonymic strength of this German pair. *Kühl* and *cool* cover a very similar semantic range in their respective languages, not just on the temperature continuum but also in the sense of ’distant, aloof’. However, *cool* also has further sub-senses which are not part of the semantic range of *kühl* – for example the sub-sense ’fashionable, great’. *Kühl* seems to be used with negative connotations more frequently than *cool* which is yet another factor that adds to the imbalance between the semantic ranges of the two lexemes.

In the case of *freezing:boiling* and *eisig:kochend*, the first and most immediately obvious difference between the two is the lack of overt relatedness between the two lexemes in German. *Boiling*, *freezing* and *kochend* are all participles used adjectivally whereas *eisig* is an adjective in its own right. The German lexeme which structurally corresponds to *freezing*, *frierend,* the participle of the verb *frieren* (’to be cold’), only covers a very small part of the semantic range of *freezing*. It can be used to indicate that something animate is cold (see example 5.7a below) but cannot, for example, be used for a body of water that is in the process of freezing over (example 5.7b) since this would indicate very strongly a personification of the lake.

(5.7) *frierend*

(a) *die* **_frierenden_** *Kinder/Tiere*
the.ART.PL being cold.NOM.PL children/animals.NOM.PL
’the children/animals who are cold’

(b) **der* **_frierende_** *See*
the.ART.SG.M being cold.NOM.SG lake.NOM.SG
’the lake which is cold’ not ’the freezing lake’

*Eisig* covers a largely identical semantic range to *freezing* apart from the literal ‘change of state’ sense of *freezing*, and is thus a relatively good match for *freezing* since the literal sense is the one which is least likely to contrast with *boiling*. The reverse of this is true for *kochend*, however; the main sense for *kochend* is the literal one (*kochendes Wasser* ‘boiling water’). *Kochend* is very rarely, if ever, used for ambient temperature (which would be *brütend (heiß)*) and has another metaphorical use, (‘boiling with anger’) which is relatively common.

While the two pairs were constructed to be similar in the two languages, the literal translation of the English pair and the difference in antonymic strength of the resulting German pair illustrate very well the vast difference between literal translation (and the use of cognates) and translation equivalents (cf. Eco 2003) which, in this case, accounts for the discrepancy in the judgement rating.

The last group of pairs on the TEMPERATURE scale is one of three asymmetrical pairs which each contain one member of the base pair in the respective language. As already mentioned above, the German pair *kalt:warm* is not included in this group. The reason for this is that the German pairs were constructed to be as similar as possible to the English ones and *cold:warm* was not in the original English questionnaire: the fact that it was an important pair on the German TEMPERATURE scale was not considered early enough.

All three English pairs, *hot:cool*, *chilly:warm* and *chilly:steaming,* are considered by the questionnaire participants to have greater antonymic strength than their German equivalents. The difference between *hot:cool* (2.68) and *heiß:kühl* (3.365) is smaller than that between the other two sets of pairs and can partly be traced back to the difference in semantic range and usage of *kühl* and *cool* which has already been discussed and may be due to fluctuations which are inherent in the methodology used to obtain the antonym judgements.

The difference in antonymic strength between the other two sets of pairs is greater than 1.0 on the Lickert scale (cf. Table 5.5) which, as has been noted above, is a significant difference and does not occur frequently in the comparison of the questionnaire data. Both the difference between *chilly:warm* and *frostig:warm* and that between *chilly:steaming* and *kühl:brütend* can be traced back to similar reasons. The German adjective *frostig* is, once again, a lexeme with a much more restricted range than the English *chilly* and is almost exclusively used in the sense of ‘cold, aloof’ (e.g. *ein frostiger Empfang* ‘a frosty welcome’), and thus the opposition between *warm* and *frostig* is restricted to contexts which would lead to lower antonymic strength. In the last pair of examples, while *chilly* and *kühl* are relatively good equivalents (better in fact than *kühl* and *cool*), *steaming* and *brütend* are largely different in semantic range: *brütend* is used exclusively for ambience temperature whereas *steaming* is also used to refer to objects of a high temperature (whether they actually give off steam or not is to a certain degree irrelevant).

Unlike the spatial opposites, there are several considerable differences between the English and German pairs on the TEMPERATURE scale. This is largely due to the fact that there is a wider range of lexemes avaiable in this lexical field and, due to their greater semantic content, it is more difficult to find translation equivalents which also match each other in the criteria for good antonymy outlined in chapter 2.5. One noteworthy trend which can be observed in the data analysed above is that the German asymmetric pairs score consistently lower than their English counterparts, but it remains to be seen whether this is also the case in the other two clusters of gradable opposites.

##### *5.2.3.2 The SIZE continuum*

The English and German adjectives on the SIZE continuum do not match in terms of semantic range, distribution and number. This was already touched on in the description of the stimulus selection above and will have to be considered more closely here. Of the eleven English pairs in the original questionnaire, only seven have German equivalents since the German pair *groß:klein* covers not only almost the complete semantic area covered by *small*, *little*, *big* and *large* (except the 'overweight' sense of *big* and *large*) but also parts of the area covered by *tall* and *short* in English (most notably HEIGHT).

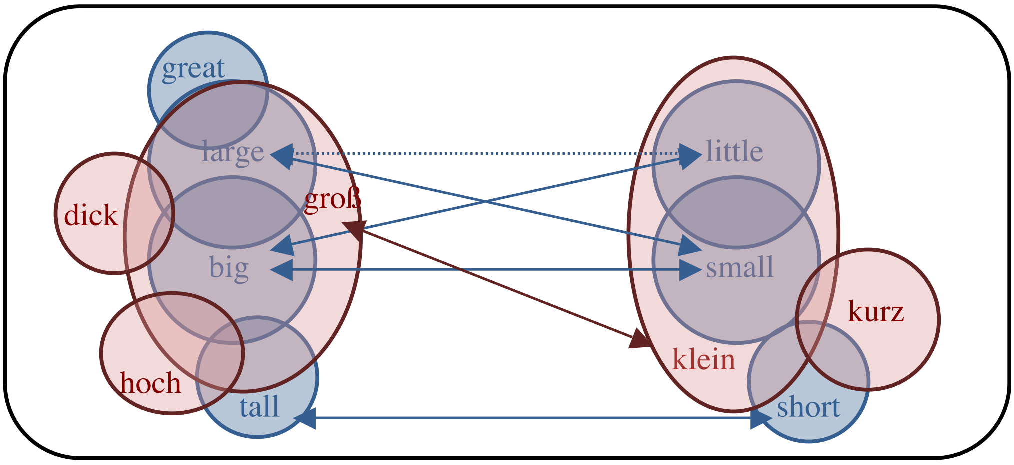



**Figure 5.1** *Overlap between the central SIZE adjectives in English and German*

Figure 5.1 above illustrates the comparison of the semantic range of the central size adjectives in German and English while Table 5.6 lists all English pairs on the SIZE continum with their German counterparts. For the most part, English and German pairs are relatively similar in terms of judgement of antonymic strength despite some considerable differences in frequency of co-occurrence as can be seen from the t-scores below.

**Table 5.6** *Cross-linguistic comparison of antonym pairs on the SIZE continuum*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| big | small | 1.225 | 10.79 | groß | klein | 1.08 | 123.709 |
| small | large | 1.24 | 25.67 | | | | |
| big | little | 1.24 | 8.13 | | | | |
| giant | dwarf | 1.45 | 1.72 | Riese | Zwerg | 1.285 | 11.824 |
| tiny | huge | 1.49 | 3.90 | winzig | riesig | 1.34 | 5.365 |
| big | tiny | 2.04 | 1.82 | groß | winzig | 2.725 | 4.564 |
| large | little | 2.05 | -3.72 | | | | |
| little | gigantic | 2.19 | 0.42 | klein | riesig | 2.8 | 5.869 |
| small | huge | 2.25 | 1.53 | klein | riesig | 2.8 | 5.869 |
| small | tall | 3.14 | 2.55 | | | | |

The slight differences in GOE-rating between the four (three excluding the much less antonymic pair *little:large*) central pairs in English and the central pair in German, *groß:klein*, can be traced back to the simple fact that the German pair is the only one available and the two members of the pair are extremely well matched in semantic range and all other criteria for good antonymy, including extremely high frequency of co-occurrence. The fact that there are several possibilities in English may contribute to the slightly weaker antonymic strength these pairs display since they are, individually, less general than the German pair (cf. Muehleisen 1997).

The other symmetrical pairs included in both languages are *tiny:huge* and *colossal:miniscule* and the nominal pair *giant:dwarf*. Both *tiny:huge* and its German equivalent *winzig:riesig* as well as the nominal pairs *giant:dwarf* and *Riese:Zwerg* are almost identical in terms of their GOE-rating scores with the German pairs being rated slightly better. The question of whether nominal pairs are generally judged to have more antonymic strength in German than in English will be considered in greater detail in 6.1.3.1. The difference between the GOE-rating of *colossal:miniscule* (1.505) and that of *kolossal:unscheinbar* (4.2) is the biggest difference of any pair between the questionnaires overall. This is largely due to the fact that the German pair was 'engineered' in a similar manner to the English one but the literal translation from *colossal* to *kolossal* does not match semantically and the opposition between *kolossal* and *unscheinbar* 'inconspicuous' in German is tenuous to say the least.

The last pairs to be discussed are the asymmetrical examples. Two of the English pairs, *small:huge* and *little:gigantic*, were translated with corresponding German pair: *klein:riesig* since there is, as previously stated, only one equivalent for *little* and *small* and *riesig* is the only common equivalent for *huge* and also for *gigantic* because it incorporates the same etymological relatedness to *Riese* as *gigantic* does to *giant*. The third English pair, *big:tiny*, corresponds to the German *groß:winzig*. Despite obvious imbalances in the t-scores of these pairs which should favour the German examples, all three English pairs score more than 0.5 higher than their German counterparts in the GOE-rating. This ties in with the data for the asymmetric pairs on the TEMPERATURE scale which also suggest that German seems to place more weight on the criterion of symmetry than English. It remains to be seen whether the results for the last cluster of gradable opposites, the adjectival pairs on the MERIT scale, also show this tendency.

*5.2.3.3 The MERIT scale*

On the merit scale, every one of the nine English pairs has a German counterpart since the size of the semantic fields in the two languages seems to be more equal than in the previous cluster. The central pairs, *good:bad* and *gut:schlecht* and *good:evil* and *gut:böse,* are very similar as regards judgements of their antonymic strength. Of the four symmetrical pairs, the ones which are located at the extremes of the scale, *excellent:atrocious* and *exzellent:miserabel*, are also very similar in terms of their GOE-rating. The second example of symmetrical pairs, *fair:poor* and *angemessen:mangelhaft*, are not quite as close in antonymic strength. This could be in part due to the translation of the English pair since *poor* would often be rendered as *schlecht* in German but since *schlecht* is part of the base pair, it was not an option here. *Mangelhaft* is very much related to academic performance and is one of the worst possible grades in the German school system. Thus its semantic range is relatively restricted compared to *poor*. Furthermore, *angemessen* is not a very fitting translation for *fair* since its meaning is closer to English *appropriate*.

The four asymmetrical English pairs and their German equivalents do not all follow the pattern observed in the previous two clusters of gradable opposites. In two cases, *excellent:bad* and *bad:satisfactory*, the corresponding German pairs (*exzellent:schlecht* and *schlecht:zufriedenstellend*) are judged to be of almost exactly identical antonymic strength. The other two pairs, however, display the pattern already highlighted in the previous discussion of asymmetrical opposites. Both *mediocre:brilliant* and *good:mediocre* are judged to be better opposites than *mittelmäßig:brilliant* and *gut:mittelmäßig*. The German pairs are reasonably well matched in semantic range and are good translation equivalents of the English pairs. Therefore, the assumption prevails that symmetry is a more critical criterion in antonym judgements in German than it is in English.

**Table 5.7** *Cross-linguistic comparison of opposite pairs on the MERIT scale*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| good | bad | 1.175 | 27.19 | gut | schlecht | 1.215 | 71.617 |
| good | evil | 1.39 | 15.77 | gut | böse | 1.265 | 68.632 |
| excellent | atrocious | 2.14 | 0 | exzellent | miserabel | 1.94 | 0 |
| excellent | bad | 2.64 | -0.47 | exzellent | schlecht | 2.625 | 1.494 |
| mediocre | brilliant | 3.61 | 0 | mittelmäßig | brilliant | 4.865 | 0 |
| good | disobedient | 3.375 | 0 | brav | ungehorsam | 2.08 | 1.698 |
| bad | satisfactory | 4.415 | 0 | schlecht | zufriedenstellend | 4.45 | 3.269 |
| fair | poor | 4.33 | 0.96 | angemesssen | mangelhaft | 4.74 | 0 |
| good | mediocre | 4.815 | 2.59 | gut | mittelmäßig | 5.62 | 9.033 |

The final pair to be discussed is not strictly speaking on the MERIT continuum but was included in the English questionnaire because it contained one of the base pairs in a different sub-sense from the one in which it was used in most of the other pairs. The German equivalent of *good:disobedient*, *brav:ungehorsam*, shows significantly greater antonymic strength in the GOE-rating than its English counterpart. The explanation for this discrepancy can, once again, be put in terms of semantic range. *Brav*, 'obedient, well-behaved' and *ungehorsam* 'disobedient' are a much better match since they cover roughly the same semantic range whereas *good* has a much wider range than *disobedient* which is as restricted in its usage as both members of the German pair. Furthermore, *disobedient* also already has a morphologically related opposite in *obedient* whereas in the German pair each member is the other's primary opposite (despite the fact that they form a triplet with *gehorsam*).

The greater importance of symmetry in German antonym judgements still receives some support from the data discussed in this section; however, it may not be as clearcut as initially assumed. Overall, most of the pairs on the merit scale, especially the two base pairs, show a relatively high degree of similarity in antonymic strength between the two languages.

*5.2.3.4 A German quartet:* einfach, schwierig, leicht & schwer

The cluster of *einfach*, *schwierig*, *leicht* and *schwer* ('easy', 'difficult', 'light/easy' and 'heavy/difficult') was chosen because it displays similarities in behaviour to the central size adjectives in English (cf. 3.4.1.3 and 4.4.2.2). There are, however, some important differences between the two clusters which will be illustrated before a discussion of the judgement ratings of the pairs formed by the adjectives below. Both *leicht* and *schwer* are polysemous and mean 'light' and 'easy' and 'heavy' and 'difficult' respectively while *einfach* and *schwierig* only have the latter meaning in each case ('easy' and 'difficult').

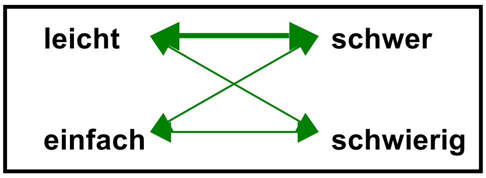



**Figure 5.2** *Antonymous relationships between* einfach*,* schwierig*,* leicht *and* schwer

Nevertheless, as can be seen from Figure 5.2, these adjectives not only form the two pairs mentioned above (*leicht:schwer* and *einfach:schwierig*) but also form pairs across the 'divide' – *leicht:schwierig* and *einfach:schwer* – which are just as acceptable. The bold arrow in Figure 5.2 signifies the double opposition between those two lexemes ('easy:difficult' and 'light:heavy') while the other three pairs are only opposed to mean 'easy:difficult'. Unlike in the case of the English size adjectives, there is no pair like *little:large* which is considered less antonymically opposed than the other pairs in the group. It seems to be a belief held by prescriptivists that the correct words to be used for 'easy' and 'difficult' are *schwierig* and *einfach* and should be used wherever possible while *schwer* and *leicht* schould be reserved for the weight of concrete objects only.[11] There are, however, a large number of constructions where *leicht* and *schwer* are the only possible collocates despite the fact that the concept is an abstract one, for example in (5.8), in which they are only used abstractly rather than to describe interaction with a heavy object.

| | | | | |
|---|---|---|---|---|
| (5.8) | *Es* | *fällt* | *mir* | *leicht/schwer.* |
| | It.NOM | fall.3SG.PRS | me.DAT | easy/difficult |
| | 'I find it easy/difficult' | | | |

The examples in (5.9) form an interesting group which serves to illustrate that, when used with a concrete noun, *schwer* cannot be used in its abstract meaning and thus cannot be used interchangeably with *schwierig*. However, in (5.9a) *schwer* is used abstractly in a fixed collocation with *erziehbar* and this phrase is very similar in meaning to the one in (5.9b) where *schwierig* is used to express a similar concept. In (5.9c), however, when *schwer* is used on its own, it can only be used in its literal, physical sense.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (5.9a) | *schwer* | *erziehbare* | *Kinder* | | | | |
| | difficult | bring up | children.NOM.PL | | | | |
| | 'problem children' | | | | | | |
| (5.9b) | *schwierige* | *Kinder* | or | *Die* | *Kinder* | *sind* | *schwierig.* |
| | difficult.PL.NOM | children.NOM | | The.ART.PL.NOM | children.NOM | be.3PL.PRES | difficult |
| | 'difficult children' | | | 'The children are difficult' | | | |
| (5.9c) | *schwere* | *Kinder* | or | *Die* | *Kinder* | *sind* | *schwer.* |
| | heavy.PL.NOM | childrenNOM | | The.ART.PL.NOM | children.NOM | be.3PL.PRES | heavy |
| | 'heavy children' | | | 'The children are heavy' | | | |

In other instances, it is possible to use *schwer* and *schwierig* and *leicht* and *einfach* interchangeably in the abstract meaning. The examples in (5.10) show that there does not need to be a difference in meaning and *schwer* and *leicht* are frequently used entirely synonymously to *schwierig* and *einfach* in both predicative and

[11] One example of such a prescriptivist stance can be found on http://www.deutschesprache-schweresprache.de/schwerschwierig

attributive use. This is very common and many speakers of German do not display any differences in usage between these partially synonymous adjectives.

| | | | |
|---|---|---|---|
| (5.10a) | *eine* | *schwere/schwierige* | *Aufgabe* |
| | a.ART.SG.F.NOM | difficult.NOM.F | task.NOM |

‘a difficult task’

| | | | | |
|---|---|---|---|---|
| (5.10b) | *Der* | *Test* | *war* | *leicht/einfach.* |
| | the.ART.SG.M.NOM | test.NOM | be.IMP.3SG | easy |

‘The test was easy.’

Nevertheless, parallel to the example of the English size adjectives, there are cases where subtle differences can be detected between the two synonymous adjective pairs. In example (5.11) below, many native speakers of German will make some sort of distinction between (5.11a) and (5.11b). Most commonly, (5.11a) is described as a decision which is more intellectual or academic and where the difficulty is deciding for or against one of two viable options whereas (5.11b) often has a strong emotional component, i.e. a decision where the outcome will have far-reaching consequences and where there is no viable alternative but the ‘right’ decision is emotionally troubling. Another way of expressing the difference is that in (5.11a) the difficulty lies in the subject matter while in (5.11b) it lies within the person making the decision.[12]

| | | | |
|---|---|---|---|
| (5.11a) | *eine* | *schwierige* | *Entscheidung* |
| | a.ART.SG.F.NOM | difficult.SG.F.NOM | decision.NOM.SG |

‘an intellectually difficult decision’

| | | | |
|---|---|---|---|
| (5.11b) | *eine* | *schwere* | *Entscheidung* |
| | a.ART.SG.F.NOM | difficult.SG.F.NOM | decision.NOM.SG |

‘an emotionally difficult decision’

This tendency for *schwer*, and also *leicht*, to refer to an emotionally difficult matter is further supported by expressions such as *auf die leichte Schulter nehmen* or *etwas leicht/schwer nehmen* (‘not to be too bothered by something’ and ‘to take something to heart’)[13]. As mentioned above, however, there are speakers who do not make a distinction at all in cases such as (5.11).

The entries from a small (55,000 word) German-English dictionary (cf. Figure 5.3 below) show the differences between the two pairs clearly: *einfach* and *schwierig* are much more restricted in their semantic range and thus their overall usage than *leicht* and *schwer,* which are polysemous but share a great deal of semantic range despite the fact that *schwer* has a greater number of polysemous meanings than *leicht* (‘severe’ (MED), ‘violent’). This strong overlap of semantic

[12] Many thanks to several German native speakers who were kind enough to answer my questions about the usage of the four adjectives and who provided many of the examples above.
[13] Literally ‘to take something on the light shoulder’ and ‘to take something light/heavy’

range between the two members of each pair and the difference between the two pairs would lead us to expect a difference in antonymic strength between the two 'base' pairs, *schwer:leicht* and *schwierig:einfach* and the other two, *schwierig:leicht* and *schwer:einfach*, as was seen in the case of *little:large* in the English example. However, the data displayed in Table 5.8 does not bear out this hypothesis.

***einfach***: simple; easy; plain; one-way (*Br* single) (*ticket*)

***leicht***: light (*fig*); easy, simple; slight, minor; TECH light (*weight*)

***schwer***: heavy; *fig* difficult, hard; GASTR strong, rich; MED *etc* serious, severe; heavy, violent (*storm etc*)

***schwierig***: difficult, hard

**Figure 5.3** *Entries for* einfach*,* leicht*,* schwer *&* schwierig *in a German-English pocket dictionary*[14]

Table 5.8 shows the GOE-ratings and t-scores for the German quartet in comparison with the English size adjectives and it can be seen that, despite marked differences in t-score among the German pairs, their GOE-rating scores are all very similar. *Leicht:schwer* is, as expected, the pair with the highest antonymic strength, but it is very closely followed by *schwierig:leicht* and *einfach:schwierig* which are judged almost identically by the participants of the GOE-rating. *Schwer:einfach* scores slightly lower but the difference between its GOE-rating of 1.34 and those of the other pairs (1.205, 1.2 and 1.13) is not significant.

**Table 5.8** *Comparison of an English and a German quartet*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| einfach | schwierig | 1.205 | 18.765 | big | small | 1.225 | 10.79 |
| schwer | einfach | 1.34 | 19.398 | large | little | 2.05 | -3.72 |
| leicht | schwer | 1.13 | 68.552 | big | little | 1.24 | 8.13 |
| schwierig | leicht | 1.2 | 9.099 | small | large | 1.24 | 25.67 |

The fact that all four pairs show almost identical antonymic strength suggests that *schwierig* and *schwer* and *einfach* and *leicht* are considered synonymous in many contexts. They also have a sufficiently large semantic area in common and display an overlap in features. The added overlap in semantic range shared by *schwer* and *leicht* makes them a slightly better pair since they are opposed in more than one sub-sense. This case is similar to the English size adjectives in terms of the number of adjectives with largely overlapping meaning but there are significant differences that lead to the different distributions in antonymic strength which can be seen in Table 5.8.

[14] Langenscheidt Pocket Dictionary German (2007), Munich/Berlin: Langenscheidt KG.

### 5.2.4 The 'gender issue' revisited

The last cluster of pairs to be discussed in this section is that with the *male:female* distinction at its centre. Twelve German pairs, two adjectival and ten nominal, which all differ only with respect to the gender/sex distinction are analysed and compared to their English equivalents. The overall distribution of the German pairs is relatively similar to that in English since the nominal pairs are roughly the same distance from the two adjectival pairs which form the base pair(s) for this cluster.

However, the first point worth mentioning when considering the GOE-rating scores in Table 5.9 below is the fact that all German pairs but one, *Schauspieler:Schauspielerin*, are judged as having greater antonymic strength than the corresponding English pairs. In one instance, that of *Mann:Frau* and *husband:wife*, this difference is easily explained as has already been discussed in 5.2.2. Nevertheless, the question why some of the other pairs, among those the base pair *männlich:weiblich*, are considered considerably better antonyms in German than in English warrants closer investigation.

**Table 5.9** *Comparison of gender pairs in English and German*

| Word 1 | Word 2 | GOE | T-Score | Word 1 | Word 2 | GOE | T-Score |
|---|---|---|---|---|---|---|---|
| male | female | 2.06 | 32.54 | männlich | weiblich | 1.23 | 51.597 |
| masculine | feminine | 1.81 | 8.94 | maskulin | feminin | 1.295 | 6.244 |
| man | woman | 2.11 | 28.67 | Mann | Frau | 1.5 | 185.818 |
| husband | wife | 3.175 | 28.01 | Mann | Frau | 1.5 | 185.818 |
| mother | father | 4.55 | 29.99 | Mutter | Vater | 3.18 | 177.123 |
| aunt | uncle | 4.525 | 10.15 | Tante | Onkel | 3.715 | 32.366 |
| mum | dad | 3.5 | 26.47 | Mama | Papa | 2.875 | 70.163 |
| nephew | niece | 3.96 | 3.60 | Neffe | Nichte | 3.78 | 10.901 |
| brother | sister | 3.415 | 17.43 | Bruder | Schwester | 2.915 | 54.757 |
| king | queen | 4.775 | 15.23 | König | Königin | 3.655 | 46.303 |
| cow | bull | 3.755 | 2.18 | Kuh | Bulle | 3.28 | 5.799 |
| actor | actress | 4.4 | 2.96 | Schauspieler | Schauspielerin | 4.465 | 6.04 |

In the German questionnaire, the adjectival pairs are both ranked among the best antonyms in the entire judgement task at 1.23 (*männlich:weiblich*) and 1.295 (*maskulin:feminin*), scores similar to those of *gut:schlecht* (1.205) or *gut:böse (1.265)* or *schwierig:einfach (1.205)* whose counterparts were all also considered excellent antonyms in English. As the two pairs are relatively similar in terms of semantic range to their English equivalents, it is not immediately obvious what the cause for this discrepancy in their antonymic strength may be. One hypothesis is the cultural difference in emphasis on gender roles between the two societies in which the languages are spoken. However, *weiblich* and *männlich* do cover a certain amount of semantic range which in English is covered by *feminine* and *masculine*, for example in cases like in (5.12) below.

| | | | | | | |
|---|---|---|---|---|---|---|
| (5.12) | *Er* | *zeigt* | *gerne* | *seine* | ***weibliche*** | *Seite.* |
| | he.NOM | show.3SG.PRS | glad.ADV | his.F.SG.ACC | feminine.F.SG.ACC | side.ACC.SG |
| | 'He likes showing his **feminine** side.' | | | | | |

If one were to adopt the same approach as in 3.4.3 and suggest that the low antonymic strength of English *male:female* is due to the fact that it does not solely encode biological sex which is, at least in everyday though if not in scientific terms, considered an excellent complementary pair, but also social gender, and thus gender roles, this should also hold true for German *männlich:weiblich.* In fact, judging by the examples above, it should have an even stronger impact on the German pair since this seems to include more of the range of *masculine:feminine* which is the pair which, in both languages, is used to denote a more gradable and variable property than biological gender. However, this is not the case and there seems to be a general tendency in this cluster towards greater antonymic strength in German than in English. This is also borne out in the analysis of the remaining pairs in this group.[15]

All nominal pairs are between 0.7 and over 1.0 different from the English pairs with the exception of *Nichte:Neffe* where the difference is a relatively slight one and *Schauspieler:Schauspielerin* which, at 4.465, is almost identical to *actor:actress* (4.4). The difference between *man:woman* and *Mann:Frau* can be explained by the fact that since the German pair *Mann:Frau* also incorporates the English pair *husband:wife* (cf. 5.1), its members will co-occur more frequently and thus the connection between the two is more firmly entrenched which makes the antonymic judgement easier and, as we saw in the English behavioural data, faster. This does not, however, account for the higher antonymic strength of the six remaining nominal pairs.

One of the biggest difference in GOE-rating score is that between *mother:father* (4.55) and *Mutter:Vater* (3.18). In 3.4.3, the low antonymic strength of *mother:father* was discussed in more detail and an attribute listing task was used to determine the difference between strong nominal pairs (e.g. *giant:dwarf*) and weaker ones (*mother:father, man:woman*). The results showed that the key attributes of the weaker nouns either did not include the underlying base pair in the list of attributes at all or it was ranked fairly low whereas in the case of the better

[15] The question whether the fact that German has grammatical gender whereas English does not influences these judgements will not be discussed here since there is no evidence in the current study to support this theory.

pairs, the member of the base pair corresponding to the noun was included as a fairly central attribute (cf. 3.4.4.2). The same attribute listing task (cf. Appendix 7) was carried out with German participants (24 participants (12 in each group)) and two examples will be used to illustrate the differences and similarities found in the data discussed above.

**ATTRIBUTE LISTING RESULTS**

Like the English task, the German attribute-listing task contained 15 nouns and the participants were asked to provide as many attributes for each item as they could think of in 60 seconds. A complete list of results can be found in Appendix 9. This section will very briefly consider the examples of *mother:father* and *man:woman*, which were already discussed in detail in 3.4.4.2, in comparison to their German counterparts.

| **MOTHER** | | **MUTTER** | |
|---|---|---|---|
| has children/gave birth | (10; 71) | ist fürsorglich | (11; 74) |
| is caring | (7; 46) | ist warm/lieb | (7; 53) |
| is loving/loves | (6; 41) | hat geboren | (5; 34) |
| is warm/gentle | (7; 49) | erzieht | (3; 19) |
| nurtures/nourishes | (5; 33) | ist weiblich | (3; 18) |
| is female/woman | (4; 27) | ist immer da | (4; 17) |
| comforts | (3; 23) | hilft/unterstützt | (3; 13) |
| works/hard-working | (3; 17) | schimpft | (2; 10) |
| cooks | (3; 12) | lacht | (2; 7) |
| listens | (2; 11) | kocht | (2; 5) |
| teaches | (2; 4) | | |

**Figure 5.4** *Attribute listing results for* MOTHER *and* MUTTER

The results for MOTHER and MUTTER are overall not very different as the lists for both categories are made up of an almost identical number of highly similar features. There are some attributes which are given more weight in German than in English and vice versa. Most notably, in German, the biological aspect of motherhood, giving birth, is secondary to the aspect of nurturing and caring for the offspring whereas in English the distribution of these two attributes is reversed. As in the English results, very few German participants found it necessary to mention ‘female’ as an attribute of mother and this feature ranks fifth in the overall list of attributes. The comparison between the second members of the pair, FATHER and VATER, displayed in Figure 5.5, clearly shows the differences in category-internal structure beween the two languages.

The German results of the listing task contain twice as many attributes as the English results presented in 3.4.4.2 and while this is still less than the number of features for MUTTER, the difference is much less pronounced in German than in English. The German category of VATER is much more similar in structure to MUTTER than FATHER is to MOTHER. While the English list of attributes limits itself mainly to attributes which refer to biological properties (‘has children’, ‘is male’, ‘is old’) and only contains one feature which refers to other properties (‘provides’), the German category is a lot richer in attributes, especially where nurturing features like ‘caring’ and ‘protective’ are concerned. These features match those listed for the category MOTHER and thus highlight the similarity of the two categories which, considering antonymy a relation of minimal difference, could be a contributing factor to the German pair’s better GOE-rating score.

| **FATHER** | |
|---|---|
| has children | (6; 43) |
| is a man/is male | (5; 38) |
| provides | (2; 15) |
| is old | (2; 14) |

| **VATER** | |
|---|---|
| ist ein Mann/männlich | (5; 35) |
| ist ein Erzeuger | (4; 31) |
| ist fürsorglich/liebevoll | (4; 27) |
| beschützt | (5; 26) |
| ist Oberhaupt der Familie | (2; 14) |
| ist ein Vorbild | (2; 14) |
| ist streng | (2; 13) |
| sorgt für Familie | (2; 7) |

**Figure 5.5** *Attribute listing results for FATHER and VATER*

The only other pairs which will be considered contrastively in this section are *man:woman* and *Mann:Frau*. Both sets of equivalents are fairly similar in their overall category structure and in the number of attributes listed by more than two participants for each of the categories. Alongside the biological characteristics, there are several important elements in both the male and the female set which are listed early on for both languages (e.g. ‘beauty’, ‘has children’ for WOMAN/FRAU and ‘human’, ‘tall’ and ‘strong’ for MAN/MANN).

The antonymic pairs of features, *male:female* and *männlich:weiblich*, are distributed in a slightly different way in German. In the results of the English listing task, which can be seen in Figures 5.6 and 5.7 alongside the German data, only two participants list ‘female’ as a feature of WOMAN whereas ‘male’ is considered the most important characteristic of the category MAN. In German this is reversed with ‘weiblich’ coming top of the list and ‘männlich’ only being mentioned once and it should thus not technically be included in the list in 5.7.

| WOMAN | |
|---|---|
| has children | (8; 42) |
| mother(hood) | (3; 22) |
| beauty/beautiful | (3; 19) |
| female | (2; 16) |
| feminine/femininity | (2; 15) |
| breasts | (2; 15) |
| nurturer/nurturing | (2; 12) |
| skin | (2; 8) |
| smells good/nice | (2; 6) |

| FRAU | |
|---|---|
| weiblich | (7; 51) |
| schön | (4; 28) |
| kann Kinder bekommen | (5; 24) |
| hat Brüste | (3; 22) |
| liebevoll | (3; 18) |
| feminin | (2; 16) |
| weich | (2; 13) |
| trägt Kleider/Röcke | (2; 13) |
| zierlich/kleiner als Mann | (2; 10) |
| Mutter | (2; 10) |
| macht Haushalt | (2; 5) |

**Figure 5.6** *Attribute listing results for* WOMAN *and* FRAU

While this discrepancy is interesting, it does not have any bearing on antonymic strength in the present case. Despite the fact that the English categories and the differences and similarities between them are closely matched by their German counterparts, the German antonym pair, as has been discussed in 5.1.3, scores much better on the GOE-ranking task.

| MAN | |
|---|---|
| male | (7; 55) |
| strong/firm | (6; 41) |
| has a penis | (5; 31) |
| not woman/opposite of woman | (4; 29) |
| human | (3; 20) |
| testosterone | (4; 16) |
| taller than women | (2; 13) |
| mankind/humanity | (2; 12) |
| has two legs | (2; 11) |
| facial hair | (2; 10) |
| has two arms | (2; 9) |

| MANN | |
|---|---|
| stark | (8; 55) |
| groß | (7; 53) |
| Frau/Gegenteil von Frau | (4; 30) |
| hat einen Bart | (3; 21) |
| ist ein Mensch/menschlich | (2; 16) |
| verdient Geld | (3; 15) |
| praktisch | (2; 11) |
| hat einen Penis | (2; 10) |
| trinkt Bier | (2; 7) |
| männlich | (1; 7) |

**Figure 5.7** *Attribute listing results for* MAN *and* MANN

The comparative data does not reveal any evidence allowing for an answer to the question of why the German pairs which are clustered around *male:female* seem to have greater antonymic strength than their English equivalents. The German data does, however, also support both the complexity and the ranking hypotheses introduced in 3.4.3. More of this data, which relates to these hypotheses but not to the *male:female* distinction, will be discussed in 6.1.3.2.

## 5.3 Summary of analysis

When the entirety of the German and English questionnaire data is taken into account, the most striking observation, in a task which is so dependent on individual judgement, is how similar most pairs are in their antonymic strength and consistent judgements between the two languages. The discussion above has focused on the cases which display the greatest difference but it is essential to emphasise again at this point that the vast majority of word pairs on the scale, even the non-antonymic control items, are within a very small margin of difference of each other.

This provides relevant evidence for two important issues, one methodological and one theoretical. On the methodological side, the question of validity discussed at the beginning of Chapter 3 can now be answered with more certainty since the level of agreement and consistency in the cross-linguistic comparison of the questionnaire data cannot be accidental. The questionnaire was not entirely true to the principles originally investigated with this type of tool, and thus its validity was not fully established by previous research. The results discussed in this chapter along with a number of other studies making use of judgement tasks, both online and offline, have shown that despite the requirement of a meta-judgement about the quality of a lexical relation the GOE-rating format serves as a successful and important investigative template.

A large number of the greater discrepancies found in the data can be traced back to mistakes or certain methodological decisions and considerations. The most frequently encountered methodological issue was the lack of an exact translation equivalent and the decision of whether to use a morphological equivalent (e.g. *frierend* for *freezing*) despite semantic differences or to use a semantic equivalent which differs in the degree of morphological relatedness to its partner (*eisig* for *freezing*). However difficult it is to tease apart the subtle differences in meaning and semantic range between some of the English and German lexemes (cf. 5.2.3), a large number of important insights can be gained and some of the differing GOE-rating scores which are the consequence of methodological decisions illustrate the importance of some of the factors which have previously not been supported by empirical evidence but have resulted from theoretical considerations.

Semantic range, for example, appears to be an influential criterion since, in cases where there is a difference between the degree of overlap of the members of the pairs in English and German, there are instances where this can be explained in terms of a lexeme's range. However, as can be seen from Muehleisen's study

(1997), determining a lexeme's semantic range can be a complicated and time-consuming process. This especially applies to a large number of the base pair adjectives since many of these can be used in a very wide variety of contexts. Therefore, the evaluation of differences in semantic range of the items discussed above, or indeed in the present research in general, is largely based on well-researched judgement rather than exhaustive corpus research. Semantic range certainly seems to have an influence on antonym judgements; however, it remains to be seen whether this is the case simply because a larger shared range allows for more opportunities of co-occurrence and thus entrenchment or whether the range allows for an easier matching process in an online concept-comparison model (cf. for example Paradis et al. 2009). The contrastive data highlights the importance of shared **semantic range** between members of an antonym pair but this factor seems to be equally important in both languages. **Purity of opposition**, **frequency of co-occurrence** and **morphological relatedness** seem to have a similar effect in German and English.

There is, however, one factor which, even after this relatively brief discussion of a small sample of the contrastive data, seems to have a greater effect on antonym judgements in German than in English. **Symmetry** has been seen to have an effect on antonym judgements (cf. 3.4.1) but no significant effect of symmetry could be seen on the reaction times in the behavioural data (cf. 4.3.1.4) in English. In German, however, asymmetrical gradable antonym pairs score much lower in the GOE-rating than their English counterparts. This is an observation which was relatively clear after a first look at the data and which was seen to be substantiated as the evidence was considered in greater detail. The analysis of the German behavioural data will show whether the greater effect symmetry seems to have on antonymic strength in German is also in evidence in this data.

If it is indeed the case that German has a different weighting of the criteria for good antonymy, this would make the internal structure of the category very similar to other culture-dependent prototype categories where, depending on context and culture, certain exemplars which have a particularly relevant characteristic are moved into the focal position or, as is the case here, from a more central to a more marginal perspective because the weighting of the individual criteria has been altered. The question of whether the contrastive analysis of the behavioural data supports this hypothesis as well as the other claims made above will be answered in the remainder of this chapter.

## 5.4 The German lexical decision task

In addition to the German questionnaire, the behavioural study discussed in Chapter 4 was also run in German, using the translation equivalents of the word pairs used in the English version and only deviating from that pattern where no equivalent was available.

### 5.4.1 Task and stimulus selection

The task used in the German behavioural study is identical to that described in 4.2 and the stimulus selection was carried out using the principles outlined in 4.2.1 based on the German version of the GOE-rating questionnaire discussed at the beginning of this chapter (5.1). The main difference lay in the presentation of the stimuli. The programme used was the same cross-modal presentation software (xmod) used for the English study and it was run on a MacBook Pro running MacOS 10.6.2. However, since there were no laboratory facilities available at the school where the experiments were conducted, the stimuli were presented via a projector onto a screen in an almost completely darkened room using the same font and font size as in the laboratory-based study. All word pairs were again capitalised to eliminate differentiation between nouns and other word classes in German, as well as for ease of reading. Each participant had an individual response box as well as a pair of headphones to shield them as much as possible from the responses of the other participants. A maximum of 14 participants were tested at once and each of them had an individual desk as well as a clear view of the projection screen. The reaction time data from this study is comparable to that from the English study and the only significant difference which has been noted is a slightly longer overall reaction time which may well be due to the difference in average item length as well as the difference in presentation. However, as can be seen from the analysis below, the differences in reaction times within the German data are very similar to those in the English data. Since the t-score data will mainly be compared intra-linguistically to the respective GOE-rating scores and only the differences between those measures in certain pairs will be compared across languages, the difference in size of the two corpora and thus the slight distortion of the t-score will not affect the analysis.

### 5.4.2 Participants

All 28 participants were native speakers of German recruited in a secondary school in Munich[16] and all were between 18 and 21 years of age. They participated in the study voluntarily and without payment but it was seen as a part of their English and German language class. Some of the German participants not only took part in the German study but also in its English counterpart (as second language learners of English).[17] Whenever a participant took part in both studies, the German task was completed first and the English second to prevent interference.

### 5.4.3 Statistical analysis

The statistical analysis is conducted in a similar way to the methods described in 4.3 but it is presented in less detail. After a brief overview of the data as a whole, the following sections focus on areas in which the German and English data show significant differences or striking similarities and these cases are then discussed more thoroughly.

Before embarking on the statistical analysis, it has to be noted that cleaning up the German data was a much more complex task than in the case of Experiment 1 (cf. 4.2.3). The error rates for the 28 subjects who participated were higher overall in this task than in either of the previous tasks. The same method of detecting bad subjects as in Experiment 1 was used but the acceptable error rate nevertheless had to be raised to 11% to allow for a large enough number of subjects. To ensure that the difference in data presentation (projector vs. individual screen) was not the cause of this marked increase in error rates, some subjects were recorded in the same perception laboratory as the participants for Experiment 1 but their data displayed the same high error rates. One possible explanation is the marginally greater average item length in German (14.06 characters vs. 11.81 characters in the English data), which generates more late responses which are coded as errors. It is in fact the case that many of the errors in the German data are not 'wrong' decisions where an opposite pair is not recognised but are 'null' responses resulting from either a late or a non-reaction. For the overall statistical analysis, almost half the subjects had to be excluded and only nine subjects in each group were deemed to

---

[16] I am very grateful to Michael Keuchel and Gina Fobbe and the members of their LK English at Wilhelm-Hausenstein-Gymnasium München for giving up valuable class time to help with my research.

[17] The data from these experiments has not been integrated into the present research since there was an insufficient number of second language learner participants.

have sufficiently low error rates. When it comes to the analysis of individual items, however, some of the subjects with higher error rates will be taken into consideration to provide an adequate amount of data.

### 5.4.4 Overview of the German data

As in the analysis section in the previous chapter, the German data will first be looked at using a succession of individual ANOVAs to determine whether there are any considerable differences between the influence certain factors seem to have in the English and German data. Not every single analysis will be discussed in detail here as this would add unnecessary bulk to the discussion and distract from the relevant analyses. Therefore, if a factor is not discussed in detail, it is implied that the results resemble those in the English data very closely. One factor to take into account when considering the statistical data in comparison is the marginally smaller sample size in German, which may have had effects on the data presentation. Therefore, the main points of comparison will be meanRT and the distribution of the reaction times over the different conditions. These factors should be relatively reliable indicators of the similarities and differences between the two datasets.

A first observation, as already mentioned above, is that the overall reaction times are slower for the German experiment than for the English one. This may well be due to the difference in the presentation of the stimuli and the higher potential for distraction from other participants. Item length is also on average about 2 characters longer in the German experiment (G: 14.06, E: 11.81) while the standard deviations are almost identical (at 4.44 (G) and 4.10 (E)). The same number of analyses was carried out in German but this section will only serve as an overview pointing out notable differences and similarities while the following sections will consider some cases in greater depth.

#### GOE-RATING AND FREQUENCY OF CO-OCCURRENCE

The two charts in Figure 5.8 below show the comparison of the mean reaction times in the different GOE-rating groupings (cf. 4.3.1.1). The English data (blue chart) displays significant differences between all three categories in the overall analysis of GOE-raing. The results of the oneway ANOVA in the German data ($F(2, 774) = 35.9552$; $p \leq 0.0001$; RSquare: 0.56) also show that all three conditions are significantly different from each other: **excellent** at 858.966, **good** at 909.774 and

**medium** at 988.255 (red chart). There were not enough instances of **poor** to warrant the inclusion of this condition in the overall analysis.

The distribution of meanRT in the analysis by antonym type is quite similar to that in English with the conditions **excellent** and **good** resulting in faster reaction times than **medium** and **poor** in the analysis of both antonyms and complementaries. The case of converses, however, differs from the English distribution as the differences in reaction times between the conditions in the category converses is not significant (p ≤ 0.7418). This reflects the distribution of nominal converses seen in the GOE-rating in both English and German but is not reflected in the English behavioural data. The fact that the pairs which in the converse group in English (e.g. *husband:wife*) obtained relatively fast reaction times despite having low GOE-ratings were rated much higher in the German GOE-data will also have played a part in these more even results.

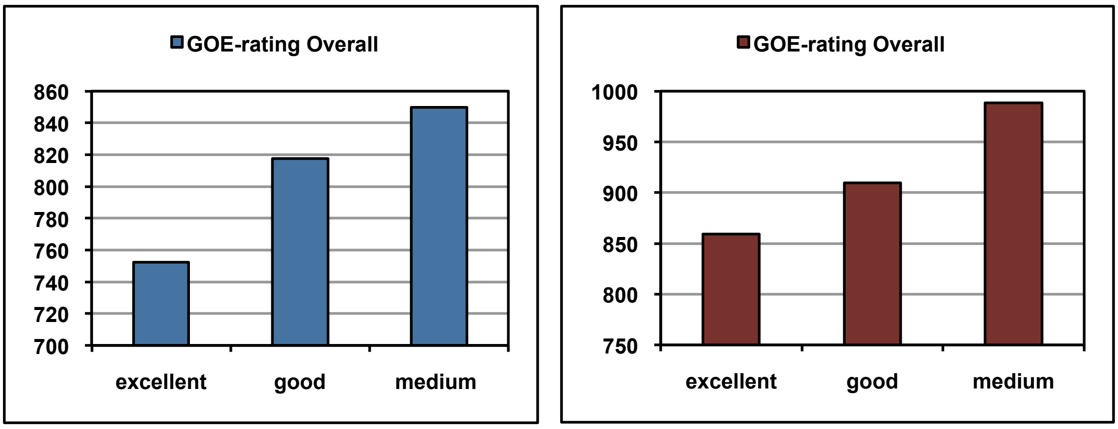



**Figure 5.8** *Comparison of MeanRT (in ms) by GOE-rating in English and German*

Frequency of co-occurrence had a substantial effect in the English behavioural data which was not always in line with the results of the judgement task. This raises the question of whether the type of task in the behavioural experiments exaggerated the effect of frequency of co-occurrence on antonym judgements because, in the case of a pair with a high t-score, the closer associative relation between the two lexemes facilitates faster recognition and thus a faster reaction regardless of the amount of antonymic strength involved. The overall results in the German data are very similar to those presented in 4.3.1.2. There is a difference in reaction time between all conditions but not all of these differences are statistically significant. In the English data, the difference between medium and low t-score was not significant whereas the German analysis (F (3, 774) = 52.6106; p ≤ 0.0001; RSquare: 0.56) shows no significance for the differences between **high** (981.298) and **medium** (1030.267) and **medium** and **low** (1124.808). **Very high** (851.500) is significantly different from all other conditions and **high** is significantly different from **low**.

In the separate analyses by antonym type, all three analyses show significant differences between at least some of the conditions. The distribution in the three categories is as expected: the higher the t-scores, the faster the reaction times. This applies to all three types of antonyms. However, there is no significant difference between the conditions **high** and **medium** in any of the three opposite types. It seems that the extremely high t-scores in the category **very high**, and the low scores which indicate that the level of co-occurrence could be attributed to chance, have the biggest influence on the reaction time data. The fact that the highest t-scores had the biggest effect was also observed in the English data. The analyses by GOE-rating are also all statistically significant; however, only the group of excellent pairs shows the 'classic' distribution in which all three conditions are significantly different (there are no pairs with the t-score rating **low** in this group). In all other groups, the only condition which differs significantly from the others is **very high**. The questions this raises regarding the influence of frequency of co-occurrence are similar to those discussed in 4.3.1.2. It seems that an extremely high frequency of co-occurrence does have a considerable influence on the speed with which a word pair is recognised as an opposite pair and, as could be seen in the discussion of the nominal pairs which obtained very fast reactions in the English data, it seems that this very strong associative relation aids recognition in general and 'muddies the water' in terms of antonymic strength since the results are heavily influenced by associative strength overall (cf. 6.1.1.5 for a detailed discussion of the influence of frequency of co-occurrence).

#### SEQUENCE, SYMMETRY AND MORPHOLOGICAL RELATEDNESS

In complete accordance with the English data, there is no statistically significant difference between the reaction times to pairs appearing in their preferred and dispreferred orders. It seems, however, that pairs which have such an ordering preference are recognised more quickly since in both the English and the German data show differences between **preferred** and **dispreferred** on the one hand and those pairs which are in the group **no order**. This could, however, also be due to the fact that many of the less conventional items included fall into that third category and those items raise the mean reaction time of that group. The data supports other studies (Paradis et al. 2009) which have stated that, although some pairs do have a distinctive preference for sequence and there are certain criteria which determine this sequence (cf. Jones 2002 and 2.2.2), this does not have an effect on their antonymic strength. However, for individual pairs it does make a significant difference and some of these pairs will be discussed in 6.1.2.3.

Symmetry is a significant factor in the overall analysis which shows, in accordance with the GOE-rating data, that symmetrical pairs were recognised significantly faster than asymmetrical ones ($F\ (1,\ 774) = 23.4779$; $p \leq 0.0001$; RSquare: 0.56). While the English analysis did not yield any significant differences in the overall analysis for symmetry, the German data shows significant differences in both the overall and the subset (gradable opposite pairs). This data seems to lend support to the hypothesis put forward in the analysis of the GOE-data that symmetry may be a more influential factor in German than in English (cf. 6.1.1.3).

In the English data, morphological relatedness did not have an overall effect on the reaction times once the long items were removed from the analysis since several of the longest items were morphologically related adjective pairs (e.g. *interested:disinterested*). The German data shows significantly faster reaction times for the unrelated items ($F\ (1,702) = 13.7086$; $p \leq 0.0002$; RSquare: 0.53), even with all items above 18 characters in length removed from the analysis.

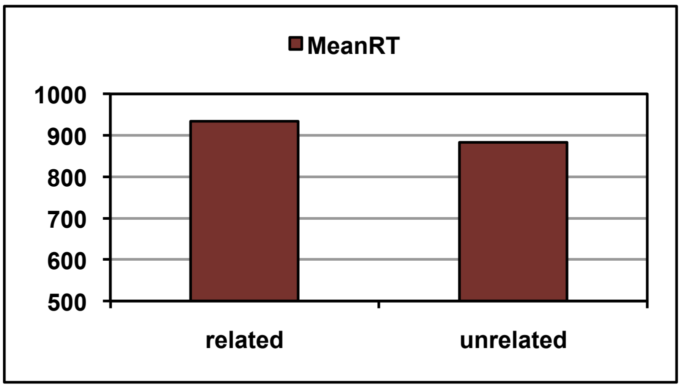



**Figure 5.9** *MeanRT (in ms) by morphological relatedness*

Figure 5.9 above shows the significant but not extremely large difference between the two conditions. There are some differences in the types of opposites included in this category in the two languages, which may have had an effect on the results above. Since this is an area which shows considerable discrepancies and morphological relatedness is an important criterion in the antonym literature, it will be discussed separately in 5.4.5.2.

#### WORD CLASS AND ANTONYM TYPE

The last two criteria to be discussed in this section are those of word class and antonym type. In the English data, both groups displayed significant differences with respect to some of the conditions under investigation. The analysis for word class in English was repeated after excluding some of the nominal pairs which obtained uncharacteristically fast reaction times due to extremely high associative strength.

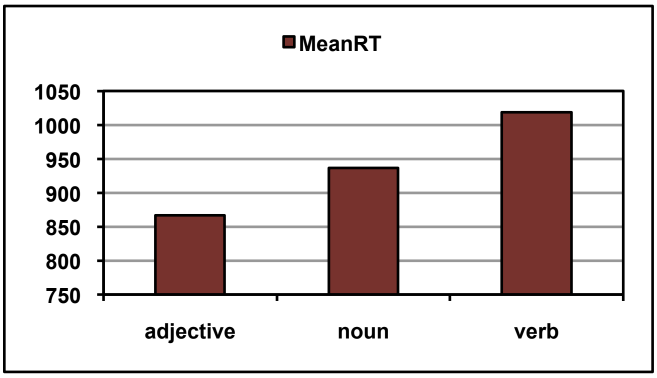



**Figure 5.10** *MeanRT (in ms) by word class*

Figure 5.10 shows the distribution by word class before taking out any of the items which were removed from the English analysis. It shows that all three conditions are significantly different from each other with adjectives being the fastest and verbs the slowest ($F\ (2, 775) = 46.1962$; $p \leq 0.0001$; RSquare: 0.56). However, using the same criteria as in the analysis of the English data, once the counterparts of even two of the pairs which were removed from the English sample (*Mama:Papa* and *Mutter:Vater*) were taken out of the dataset, the difference between verbal and nominal pairs was no longer significant while both conditions still resulted in significantly slower times than adjectival pairs. There is one other pair which scored higher than would have been expected and this may well have been due to the situation the participants were in: *Schüler:Lehrer* ('pupil:teacher') obtained a faster reaction time than was expected from both GOE-rating and t-score. The reason for this could be the fact that the experiment was carried out in a school and that all participants were pupils. This pair, however, remained in the analysis.

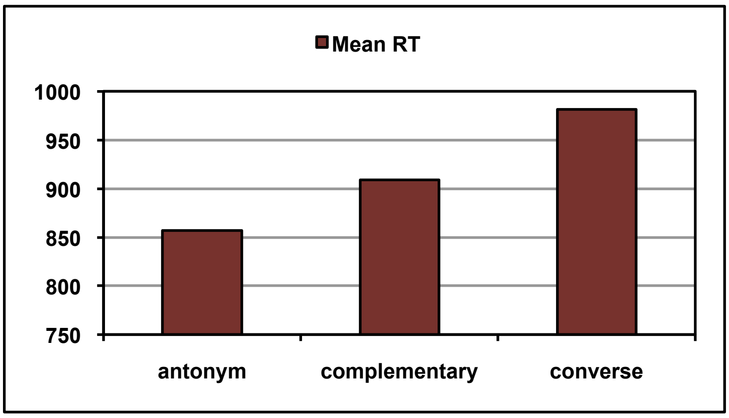



**Figure 5.11** *MeanRT (in ms) by antonym type*

While the results of the English analysis by antonym type only shows a difference between **antonyms** and **complementaries** on the one hand and **converses** on the other, the German data shows significant differences between all three antonym types. There are some instances in the German data where items received very few responses and the mean reaction time is thus not very representative.

Figure 5.11 shows the distribution across the three conditions of the analysis carried out on the whole dataset ($F(2, 775) = 32.4671$; $p \leq 0.0001$; RSquare: 0.56). The difference between converses and complementaries is greater than that between antonyms and complementaries but this is simply a tendency and no conclusions can be drawn from it. When some of the items which resulted in few correct responses were taken out, there was a slight decrease in the mean reaction time of the complementaries but the difference between the conditions **antonym** and **complementary** was still significant. The discrepancies between the German and English data discussed in this section will be addressed either below (cf. 5.4.5.2 for morphological relatedness) or in Chapter 6, where each of the criteria will be analysed separately taking into account all data gathered for this study and relating it to the theoretical background of the discussion of antonym canonicity.

### 5.4.5 A closer look – qualitative discussion of three cases

The previous section gave a brief overview of the German data overall in relation to the English results which were presented in 4.3.1. There were several instances where the two sets of data showed considerable similarities and other cases in which the results were remarkably different. While the aim of the previous section was to provide an overview of the whole dataset, this section will consider three separate cases in greater detail. First of all, two of the gradable adjectival clusters, SIZE and TEMPERATURE, which were discussed earlier in this chapter (cf. 5.2.3) from the perspective of the judgement task, are revisited and the behavioural results for both are compared. The second section takes a closer look at morphological relatedness and attempts to shed light on the conflicting results in the English and German data. Lastly, the results of the German quartet introduced in 5.2.3.4 are investigated and compared to the results of the English central SIZE adjectives from Experiments 1 and 2.

*5.4.5.1 TEMPERATURE and SIZE – German vs. English*

The discussion of the GOE-rating results of the nine pairs on the TEMPERATURE continuum which were discussed in 5.2.3.1 showed that the results for the German pairs differed considerably from those of their English counterparts. The only agreement between the data was on the pairs on the extremes of the GOE-rating scale – *hot:cold* (*heiß:kalt*), which was judged to be an excellent opposite pair, and *cold:mild* (*kalt:mild*), which was a rather poor example.

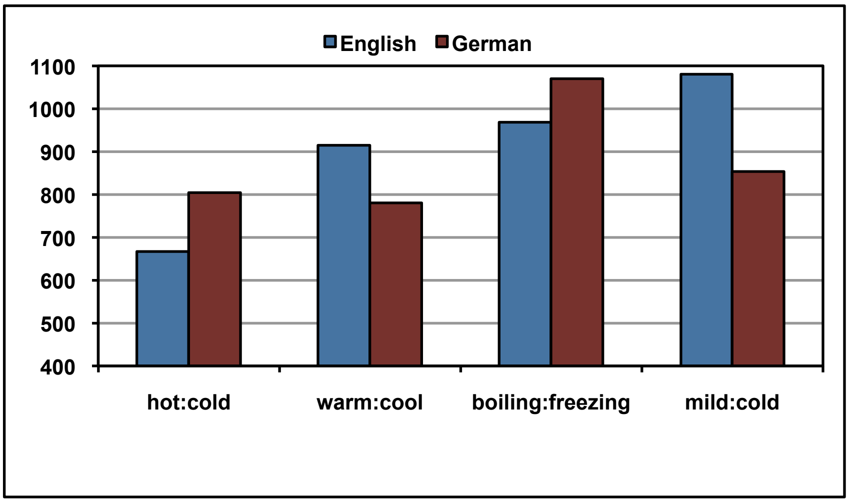



**Figure 5.12** *Comparison of meanRT (in ms) for pairs on the TEMPERATURE scale*

It can be seen in Figure 5.12 above that the English behavioural data for those pairs supports the results of the GOE-rating since those pairs are in similar positions on the reaction time scale. The English results, in fact, match those of the GOE-rating very well and the four pairs displayed below are in the same order as they are in the judgement task. The German results, on the other hand, do not match the GOE-results as well. One result which is unsurprising is the slow reaction time for *kochend:eisig* (the equivalent of *boiling:freezing*) since it has already been established that German does not have a pair of lexemes which share enough semantic range and/or are sufficiently conventionalised to cover the extremes of the TEMPERATURE scale. While *eisig* is an excellent translation of *freezing* in many circumstances and *kochend* is the literal translation for *boiling*, there is not enough overlap in semantic range to make these two lexemes a good antonym pair. This was noted in the discussion of the results of the judgement task and also manifests itself in the behavioural data. The reaction times for the other two pairs are both significantly faster than those for their English counterparts. *Warm:kühl* has a reasonably high t-score which could have had an influence on the results but the case of *mild:kalt* is different. The fact that this pair has a t-score in the medium range in German (rather than in the low range like its English counterpart), may have had an influence but there were only very few correct responses for *mild:kalt* which suggests that some resonses were so slow that they were not counted in the data. If those instances had been counted, this pair would have fared a lot worse in terms of meanRT and the data would have provided a more accurate account of its antonymic strength. This case shows that some of the differences seen on this level of analysis can be attributed to the selection of the stimuli on the one hand and on the experimental set-up on the other. These are, however, isolated examples which do not affect the valitdity of the data overall but have a greater effect in an analysis at the micro-level.

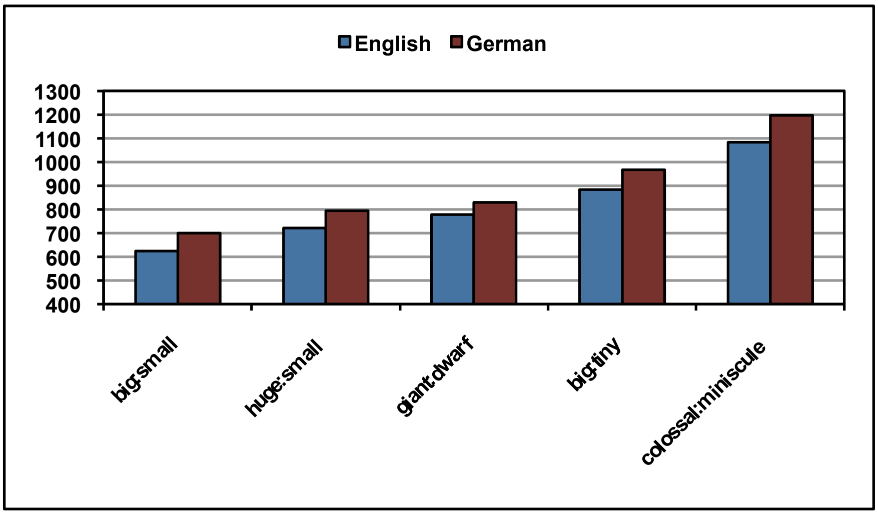



**Figure 5.13** *Comparison of meanRT (in ms) for pairs on the* SIZE *continuum*

In the case of the SIZE continuum, the differences in GOE-ratings between the English and the German pairs were less marked. The only differences which were observed were the lower ratings for all asymmetrical German pairs (cf. 5.2.3.2) and the big difference between *colossal:miniscule* and the rather ill-matched German pair *kolossal:unscheinbar*. For the behavioural task, the German pair was changed to *kolossal:mini* but it is by no means certain that this is a better combination although it possesses slightly more antonymic strength than the original pair according to some German native speakers.

The comparison of the results in Figure 5.13 show that the mean reaction times for all German pairs are very similar to those of the English pairs. In fact, when it is taken into consideration that the German reaction times are on average slightly longer than the English ones, the distribution above can be considered identical. In terms of frequency of co-occurrence, the pairs were reasonably well matched in both languages apart from the nominal pair *Riese:Zwerg* ('giant:dwarf'), which had a considerably higher t-score in German than in English (G: 11.842, E: 1.72). The example above shows clearly that, despite the differences which have been discussed and, in many cases, highlighted and used as starting points for several of the analyses, there are considerable similarities in both the GOE-rating and the behavioural data which cannot and should not be overlooked since these similarities can be used to support important claims about the nature of antonymy as a lexical relation (cf. 5.5).

The difference in GOE-rating for asymmetrical pairs which was found in the contrastive analysis of both the SIZE and TEMPERATURE fields is supported further by the behavioural data presented above. German not only shows a significant difference of symmetry in the sub-set of gradable antonyms but also in the analysis of the whole dataset. This lends further credibility to the hypothesis that symmetry is more influential in antonym judgements in German and that there are differences in the weighting of certain criteria across languages.

*5.4.5.2 Morphological relatedness – a contrastive perspective*

When morphological relatedness is briefly considered contrastively in the discussion of the GOE-rating analysis in 5.2.1, no clear patterns of either differences or similarities emerge. There are cases in which the morphologically related pair scores higher than its unrelated equivalent in the other language (*König:Königin* and *king:queen*) and other cases where there is no significant difference between the related and unrelated pairs (*buy:sell* and *kaufen:verkaufen*). One result which is the same in the English and German data in both the judgement task and the behavioural data is that the 'best' opposites are always morphologically unrelated, which is in keeping with the results of other work on opposition (cf. also Gross et al. 1989, Charles et al. 1994, Paradis et al. 2007 & 2009). It is also the case that, in all data, the worst opposites are also unrelated when those items which are extremely long are not taken into consideration. Thus, the overall distribution of related and unrelated items is fairly similar in both GOE-rating data and in the data gathered in the behavioural study.

However, as already mentioned above, the ANOVA performed on the German data results in a significant difference of meanRT in favour of the unrelated items in the study while the analysis of the English data does not show any significant difference between the two conditions. However, whether this can be taken as an indication that morphological relatedness has a different weighting as a criterion for good antonymy in the two languages will have to be established by taking a closer look at the related pairs in the German study and also considering the results obtained for related and unrelated pairs in Experiment 2 (cf. 4.4.2).

There are 13 morphologically related pairs in the German behavioural task (compared to eleven in the English task) and these include pairs from all three word classes (six adjectival pairs, four verbal pairs and three nominal pairs) and all three types of opposition (five antonyms, six complementaries and two converses). The morphologically related items in the English data are, apart from two related nominal paris, one converse (*employer:employee*) and one complementary (*actor:actress*), all adjectival and mostly gradable. This difference in distribution between the English and German stimuli may already provide part of the explanation of the slower overall reaction times in German since verbs and nouns were found to have slower reaction times overall in both languages.

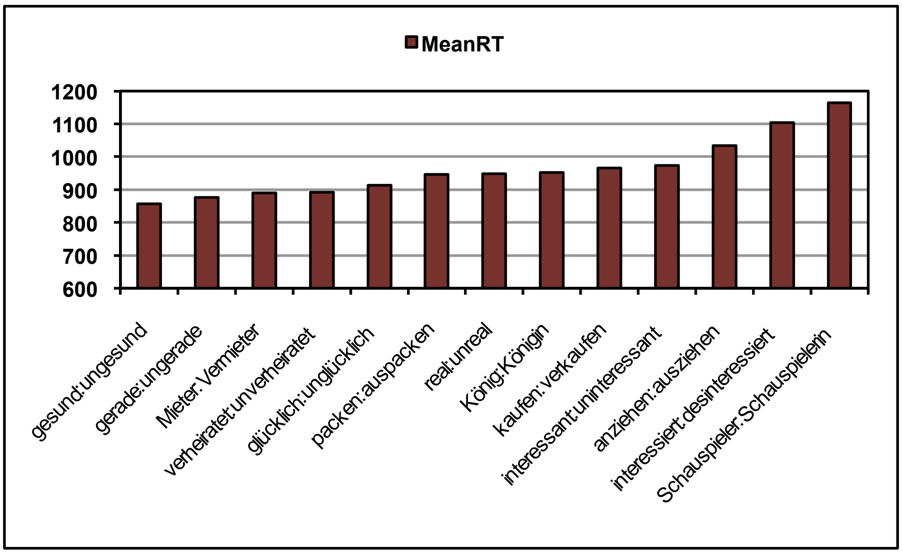



**Figure 5.14** *MeanRT (in ms) for all morphologically related pairs in German*

One question which may be asked following the discussion of repetition priming in 4.4.2 is why this does not seem to apply in the task under discussion here since the related pairs should be able to benefit from the repetition of a significant amount of graphemic material. It is clear from the chart above (Figure 5.14) that item length plays a role in these judgements but is not the sole factor determining the distribution. It may be that the simultaneous presentation of the two members of the pair does not allow for repetition, or indeed any kind of significant priming effect.

Despite the fact that the German analysis does show a significant effect, the conclusion can be drawn that morphological relatedness does not have a great effect on the antonym canonicity of a pair. Other factors seem more dominant, especially when the pair is not an adjectival one. There are, for example, a large number of pairs in the GENDER cluster which are morphologically related by suffixation. However, these pairs are not generally considered very good examples of opposition since the salience of their opposition is weaker than that of other pairs and the relatedness does not add to their antonymic strength. The related pairs which are most strongly related are those in which the prefix encodes a simple negation of the property encoded by the base lexeme (e.g. *happy:unhappy*). Since these pairs are truly minimally different, in form as well as meaning, they are consistently rated high in judgement tasks. This may also, in part, apply to those lexemes where the affix encodes directional opposition (*pack:unpack* or *dress:undress*). It seems that with increasing complexity of the relationship encoded by the affix, antonymic strength declines (e.g. German *ver*-). An investigation focusing on morphological relatedness would require great care to be taken in its experimental design in order to avoid confounding effects of repetition priming and the added complication of increased item length which does not affect many unrelated pairs.

#### *5.4.5.3 The German and English quartets compared*

In analogy to the analysis of the relationships between the central size adjectives in English, 5.2.3.4 focussed on a similar German quartet: *schwierig, schwer, einfach* and *leicht*. Unlike in the English case where the primary meaning of all four adjectives is related to the concept of SIZE, the four adjectives in the German cluster span two concepts: WEIGHT and DIFFICULTY. A reasonably thorough analysis of the meanings of the four adjectives in German and their areas of overlap has already been carried out and will only be referred to here. As was the case in the English example, both the GOE-rating and the experimental evidence showed clearly that there were relationships of varying strengths among the four words.

The German GOE-rating results for the four possible pairs were considerably closer together than in the English case where *large:little* was rated much lower than the other pairs. *Schwer:einfach* was the pair which displayed the lowest GOE-rating at 1.34, which however was only 0.14 lower than that of *einfach:schwierig*. The original hypothesis that *schwer:leicht* should score highest, since it incorporated opposition in both concepts, was confirmed by the GOE-rating results but the assumption that *schwierig:einfach* should score higher since the lexemes share a greater amount of semantic range was not confirmed.

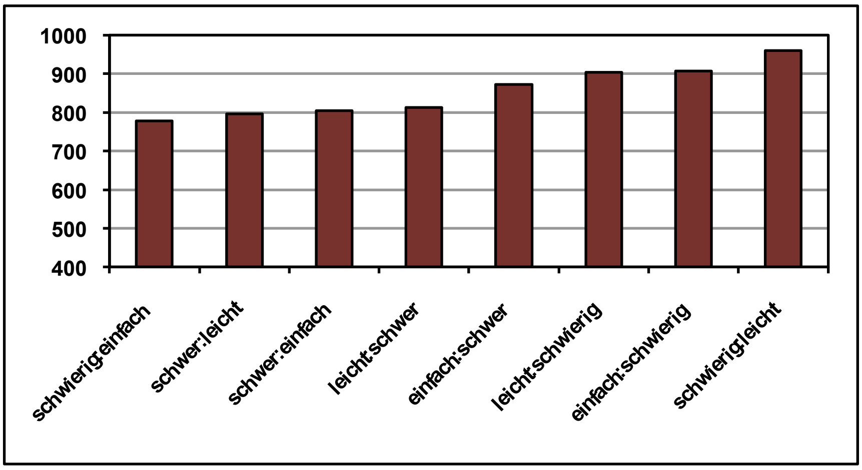



**Figure 5.15** *MeanRT (in ms) for German quartet*

Although the four German pairs display considerable differences in t-score (see Table 5.10 below), their GOE-ratings are remarkably similar. Figure 5.15 shows the mean reaction times for all eight possible pairings ordered by reaction time only. The first four pairs are all very close in reaction time whereas the other pairings are not recognised as opposites quite as fast and one pair, *schwierig:leicht*, has a considerably slower mean reaction time than the others (960.75).

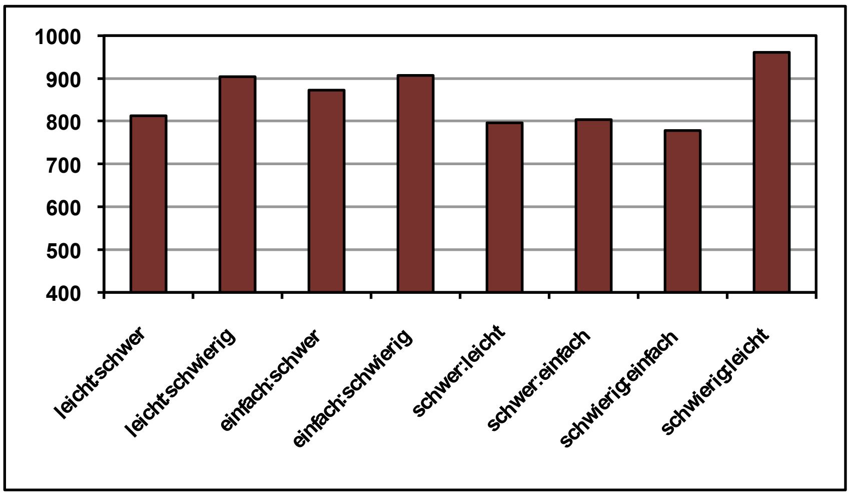



**Figure 5.16** *MeanRT (in ms) of German quartet by base adjective*

Figure 5.16 directly compares the mean reaction times for each adjective with both possible partners. In two cases, those of the base pairs *schwer* and *einfach*, the difference between the two possible pairings is minimal, whereas in the other two there is a difference, of about 100ms in the case of *leicht* and nearly 200ms for *schwierig*. It seems that *schwer* and *einfach* do not show a particular preference for either partner (cf. also Table 5.10 below). *Schwierig* especially seems to make a much better pair in combination with *einfach* than with *leicht* and the same is the case, but to a lesser extent, for the reverse pairing *leicht:schwierig*.

**Table 5.10** *Comparison of MeanRT and GOE-rating scores for the German quartet*

| BASE | ANT 1 | MEAN RT | GOE | ANT 2 | MEAN RT | GOE |
|---|---|---|---|---|---|---|
| leicht | schwer | 813.333 | 1.13 | schwierig | 904.142 | 1.2 |
| einfach | schwer | 872.333 | 1.28 | schwierig | 907.222 | 1.18 |
| schwer | leicht | 796 | 1.13 | einfach | 803.875 | 1.4 |
| schwierig | leicht | 960.75 | 1.15 | einfach | 777.666 | 1.23 |

This discrepancy is noteworthy especially since it is not evident from the GOE-rating data, unlike the difference in meanRT between *large:little* and the other three pairs of size adjectives. A similar discrepancy was observed in the data gathered in Experiment 2 (cf. 4.4.2.2) where *big:little*, a pair which was rated exactly the same in the GOE-rating task as *large:small*, displayed a much longer reaction time than the judgement task results would have predicted (cf. also Table 5.11). Since Experiment 2 was a priming task, this was attributed to inhibitory priming since the presentation of *big* activated *small* and not *little* and a readjustment was thus necessary when *little* was displayed instead of the expected *small*.

**Table 5.11** *Comparison between the German and English quartets*

| Word 1 | Word 2 | MeanRT | GOE | T-Score | Word 1 | Word 2 | MeanRT | GOE | T-Score |
|---|---|---|---|---|---|---|---|---|---|
| einfach | schwierig | 842.444 | 1.205 | 18.765 | big | small | 624.69 | 1.225 | 10.79 |
| schwer | einfach | 838.104 | 1.34 | 19.398 | large | little | 884.625 | 2.05 | -3.72 |
| leicht | schwer | 804.666 | 1.13 | 68.552 | big | little | 806.41 | 1.24 | 8.13 |
| schwierig | leicht | 934.333 | 1.2 | 9.099 | small | large | 648.285 | 1.24 | 25.67 |

Table 5.11 compares the results of the German quartet with those of the English central size adjectives discussed in 4.4.2.2. However, since the data used for that discussion was obtained in a separate task constructed specifically to bring out the differences between the pairings, it is not comparable to the data obtained in the German experiment. Therefore, the reaction times used for the comparison are those from the English Experiment 1. It is clear from the English data above that this inhibitory effect is also to some extent present in Experiment 1 where both members of the pair are presented simultaneously. *Big:little* displays a reaction time which is much closer to that of *large:little* than that of the other two pairs despite the fact that it is usually considered one of the two canonical pairs alongside *large:small*.

A similar conclusion can be drawn for the German pairs: *schwer:leicht* is, for reasons which have been discussed above, the pair which displays the fastest overall reaction times as well as the closest match between the times of the two possible combinations. *Schwierig:einfach* and *schwer:einfach* are virtually identical in terms of meanRT and although *einfach* seems to have a slight preference for *schwer* over *schwierig* this is not significant and no conclusions can be drawn. The last pair, *schwierig:leicht*, not only shows the greatest discrepancies between base adjective and partners in both members of the pair but also obtains a significantly lower reaction time than that of the other three pairs. This could also be attributed to an inhibitory effect since, as the GOE-rating data shows, it is by no means considered a ‘bad’ opposite pair. It would be interesting to see whether a priming task such as the one conducted in Experiment 2 would confirm this hypothesis if it were conducted for these German pairs.

### 5.4.6 Summary

The conclusions which can be drawn from the German behavioural data and its comparison with the data from the English Experiment 1 can be separated into methodological insights and the consequences for antonym canonicity and the factors which influence antonymic strength. There is, of course, a strong relationship between these two sets of conclusions, since the methodological realisation influences the data and its impact and reach, but the consequences of the behavioural task and the insights gained from the data will be discussed in two further sections: the first deals with the methodological aspects of the study while the second and more important discussion centres on the question of what the data contributes to the already exisiting body of research and how it compares to the results discussed in Chapter 4.

From a methodological perspective, there were several noticeable differences between the English data from Experiment 1 and the German data. The most striking of these were the much higher error rates in the German data as well as the longer reaction times overall. As already mentioned in the discussion of results, the longer reaction times may be related to the fact that the German experiment was conducted in a classroom and the participants did not have individual screens to complete the tasks and although they were wearing headphones to minimise distraction, they will have been influenced by their surroundings more than the English subjects who performed the task in a purpose-built laboratory. The other contributing factor is without doubt the increased item length in the German experiment: the English stimuli have an average length of 11.81 characters while the German targets are, on average, 14.06 characters in length. Both of these factors may have contributed to the high error rate; the first since it is assumed that the subjects will also make more mistakes if there is greater distraction through the environment but also, in combination with the second factor, through the longer reaction times, since the responses are only recorded up to a certain time threshold after which they are coded as null responses by the recording device. Both factors will have played a role in the higher error rates which resulted in the exclusion of many subjects and a relatively small dataset for the analyses above. However, since the data was intended for comparative and supporting purposes, this is not a major shortcoming.

A methodological similarity which was seen in the behavioural data was the presence of slight priming effects in both tasks with simultaneous presentation of the two members of a pair which can be seen in the analyses of the quartets (cf. 4.4.2.2 and 5.4.5.3). This effect is much more substantial in the delayed presentation paradigm used in Experiment 2 and thus has a much bigger influence on the overall analysis. Since at least one of the few other studies which used lexical decision methods to investigate antonyms, Charles et al. 1994, conducted the experiment with a 350ms delay between the first and second member of the pair, this raises the question whether the resulting priming effect has distorted the results since associative priming will, of course, also have had an effect, especially in the pairs which are particularly strongly associatively related; this effect may well have exaggerated the differences between that category and the remainder of the member of the category of lexical opposition.

As far as the analysis of the German data is concerned, it shows that the factors which emerge as significant in their influence on the independent variable are very similar to those in the English data discussed in the previous chapter. **GOE-rating**, **word class**, **t-score, symmetry** and **antonym type** are all found to have a significant effect on reaction time, while the difference between the conditions in the category **sequence**, as in the English data, is not statistically significant. **Morphological relatedness** has already been discussed in greater detail above and seems to have an impact on reaction times but the reasons for this are difficult to determine.

The direct comparison of the data to the English results shows striking similarities which lend support to the similarities observed in the GOE-rating data in the first half of this chapter (cf. also 5.5) and which cannot be chance phenomena. One particular point which can be seen in both datasets is the considerable influence a very high t-score has on the independent variable. The German data showed, once again, several nominal pairs which are at best considered to have mid-range antonymic strength in the judgement task to be among the pairs which obtain the fastest reaction times. These pairs, which were the same pairs in German and English, had to be removed from the datasets for individual analyses since their extremely fast reaction time led to an exaggeratedly low mean reaction time for nouns in the analysis of word class and for the category of medium GOE-rating in the analysis of antonym types. This further supports the hypothesis that there are some pairs which are mediated through associative strength and, in a task which crucially depends on speed, this effect aids the recognition and assessment of pairs with great associative strength (cf. 6.2.2).

## 5.5 Conclusions

In the introduction to the contrastive analysis (cf. 5), the benefits comparing the analysis of the English and German were highlighted. The hypothesis that a larger pool of data would lend more credibility to the results as well as enabling a better understanding of the psycholinguistic mechanisms at work has been confirmed. The latter applies to the behavioural task in particular since a particular configuration of data may have been either discarded as an idiosyncrasy despite being a regular, if somewhat unusual, feature of the phenomenon under investigation whereas, on the other hand, it would have been possible to draw erroneous conclusions from data which in fact resulted from a weakness in the experimental design. The greater amount of data as well as the ability to compare and contrast two language systems facilitate the confirmation or discarding of hypotheses.

From a methodological standpoint, both methods used provided useful insights into the structure and nature of the category of antonymy. The judgement task proved to be a very useful and reliable tool for the assessment of the overall antonymic strength of each pair, and cases of mismatch of the results between English and German provided very useful starting points for analyses as well as an indication of methodological shortcomings of the task (especially in terms of stimulus selection cf. 5.1.3). As discussed in 5.4.6, the addition of the German results of the behavioural tasks to the results for the English data provided supporting evidence for some of the phenomena observed in the English analysis in chapter 4 both from a theoretical and a methodological perspective.

This section will bring together the evidence from the contrastive analyses of both methods, and results which have not previously been discussed will be briefly compared. The question of whether the combination of evidence from such different methods (i.e. judgement tasks and behavioural data) creates problems in terms of analysis and explanatory power will not be discussed here but will be considered in 7.2, since the use of a variety of methods is one of the ways in which the present work sets itself apart and this issue warrants thorough discussion.

The analyses conducted in the previous and present chapters show a number of discrepancies and similarities in the data which serve as starting points for analyses. Those which were considered of particular interest have been discussed separately. The present discussion focuses on two main issues: firstly whether the critera for good opposition proposed at the outset of the present work apply only to opposition in English or whether the same factors also influence antonymic strength in German as well. And secondly, it investigated what the contrastive analysis has been able to contribute to the theoretical debate about the type of relationship between antonymous lexemes. Both of these issues will be addressed in turn.

It has emerged very clearly from the analyses and discussions above that most factors which influence the antonymic strength of English antonyms have the same degree of impact on German antonyms too. One criterion which was found to be significant in the discussion of the GOE-rating data was that of semantic range (which encompasses Cruse's criterion of match of non-propositional meaning (cf. Cruse 1986)). This criterion is not part of the discussion of the behavioural results since establishing the degree of shared semantic range of two lexemes is an extremely time-consuming process and very often subjective judgements supported by lexicographical or corpus evidence prove sufficient. A slight mismatch in the degree of semantic range between the English and German pairs has been seen to result in differences in the judgement task results. Pairings which were, in addition, less conventionalised but where no equivalently conventionalised pairing was available in German proved to be judged to have weaker antonymic strength.

While semantic range was an equally important factor in both languages, there were two instances where considerable differences could be seen between the influence a certain factor had on antonym pairs in the two languages. In the judgement task, symmetry seemed to be a more crucial criterion in German than in English (cf. 5.2.3) and all asymmetrical German pairs were rated considerably lower than their English counterparts. The behavioural data provided further evidence to support this claim. However, to determine whether this criterion does in fact have different degrees of influence cross-linguistically, a specific task would have to be designed which would control carefully for all other factors which could have influenced the data.

The second idiosyncrasy, which was initially detected in the GOE-rating data was that nominal pairs seemed to be rated to have higher antonymic strength in German than in English (cf. 5.2.2). This was confirmed to a certain extent by the behavioural data in the analysis of the distribution of mean reaction times for the criterion **word class**. In the English data (cf. 4.3.1.7), nouns are recognised slowest overall while in German, even after a re-analysis following the exclusion of two particularly high-scoring pairs (*Mama:Papa* and *Mutter:Vater*), verbs are slower to be recognised than nouns. The difference is, in the second analysis, no longer statistically significant but there is still a considerable difference in reaction time between nouns and verbs. This may be tied to a difference in preference in syntactic structure between the two languages with German having a greater tendency to give preference to nominal constructions while English prefers verbal expressions.

Lastly, and in my opinion most importantly, the cross-linguistic comparison of lexical opposition has resulted in remarkable similarities in all results, which cannot be coincidental. Thus there must be at least a conceptual basis for lexical opposition, a conclusion which stands in contrast to the position of the 'lexical categorical model' (e.g. Gross & Miller 1990) supported by many earlier approaches to antonymy which claims that lexical opposition is a relation between word forms and not concepts. However, if the same concepts result in the best representatives of the category of opposition in different language, there has to be a conceptual aspect involved at least to some extent in antonym canonicity. The data discussed in the last three chapters is in accordance with the cognitive approach taken in several recent approaches and in the present work, which suggests that antonym judgements are made on the basis of conceptual similarity and opposition but that the recognition and processing of highly conventionalised pairs at the top of the scale may well be aided by the entrenchment provided by frequent lexical association. The use of comparative data thus provides evidence which brings us closer to answering the 'chicken-and-egg' question of whether frequency of co-occurrence is a cause or a consequence of high antonymic strength. This point, among others, will be discussed further in the following chapter which relates the findings which were presented and discussed in the past three chapters to the theoretical concepts which are said to underpin the relation of lexical opposition.

# 6. Antonymy canonicity – revisited

The opposite of *less* is *more.*<br>
What's better? Which one are you for?<br>
My question may seem simple, but<br>
The catch is – more or less of *what*?<br>
[...]<br>
The best thing's to avoid excess.<br>
Try to be temperate, more or less.<br>
(Wilbur 2004: 534)

The discussion and analysis of the different strands of data has provided substantial evidence for many of the factors which contribute to antonymic strength. Some of these criteria were discussed in the previous chapters with regard to a specific dataset (e.g. GOE-rating results) but the overall theoretical implications themselves have not yet been considered. The first section of this chapter will consider the criteria proposed in 2.5 and all other factors which have been proposed throughout the data analysis from a theoretical perspective and conclude with the proposal of a model based on the conceptual structure and mechanisms involved in antonym canonicity.

In the second section (6.2) an answer is proposed to the first research question raised in the introduction (cf. 1.3). Does the data collected by different empirical means support the hypotheses proposed throughout the present work? The answer is based on the results which can be seen when the three strands of data which have been collected are combined and the differences and similarities between the results are investigated closely. The discussion incorporates an evaluation and introduction of the only other multi-method approach to the study of antonymy to assess whether the evidence presented in the data chapters of this study matches the conclusions drawn by Paradis et al. (2009) or whether the broader approach of this study as well as the differences in empirical methods lead to different results (cf. 6.3).

## 6.1 What makes an antonym pair canonical?

Throughout this work canonicity has been understood as the degree to which an antonym pair is considered antonymous, and it was thus very closely linked to antonymic strength, the two terms often being used synonymously here. The notion of a continuum of canonicity was introduced in 2.6 and is discussed further in this chapter. One definition which encapsulates the two key components of ‘good antonymy’ – which in turn are dependent on a number of factors – is that by Murphy (2003: 31) who states that ‘[c]anonicity is the extent to which antonyms are both semantically related and conventionalized as pairs in language.’ The empirical methods used in this study were intended to aid the investigation into which factors influence the canonicity of opposite pairs. In 2.5, a set of criteria for canonical antonymy consisting of factors collected from several separate investigations of antonymy (Cruse 1986, Muehleisen 1997, Murphy 2003, Jones 2002/2007, Paradis et al. 2009) was put forward and these criteria were used to structure the analysis of the different types of empirical data. They were found to have different degrees of explanatory power for the phenomena observed in the data. On several occasions, these criteria proved to be insufficient to explain the empirical results. Therefore, new criteria – or hypotheses as to what could constitute new criteria – were introduced in the data analysis to attempt to shed light on the differences observed in the material.

The question of what makes an antonym pair ‘canonical’ as well as a discussion of the concept of antonym canonicity itself are the focus of the following sections. The ‘old’ criteria proposed in 2.5 are re-evaluated as to their usefulness and predictive strength in light of the evidence collected and presented in Chapters 3, 4 and 5. Some additional criteria – for example inherent binarity and morphological relatedness – are discussed in the next section despite the fact that did not constitute the main focus of the investigation (cf. 6.1.2). Lastly, a new approach to account for the different degrees of relatedness between the two members of an antonym pair is presented (6.1.3) and the proposal of a holistic, cognitive model which incorporates the criteria discussed in the preceding sections is addressed.

### 6.1.1 Old criteria

This section will focus on the ‘old’ criteria which have been proposed to account for the differences in antonymic strength of antonym pairs. Each of the five factors will be discussed in turn and insights gained from the more detailed analyses in the previous chapters are compiled and related back to the theoretical foundations introduced in Chapters 2 and 4.1.

*6.1.1.1 Minimal and sufficient difference*

The fact that the relation of opposition is dependent on minimal difference has been considered one of the most important factors in determining which antonyms are judged to be good examples of the relation. It has become apparent throughout the present research that the more definable and containable the differences between the meanings of the two members of an antonym pair are, the more antonymic strength this pair will display. However, the criterion of minimal difference only seems to apply to the concepts and not to the lexical items which encode them. Lyons (1977) goes so far as to say that most common opposites are morphologically unrelated since that shows the dichotomy involved in antonymy better as the two lexical items then have nothing in common. This cannot be considered completely accurate since there are several excellent examples of opposition (e.g. *include:exclude* and *incaccurate:accurate*) in which both members of the pair share a large amount of lexical material. Nevertheless, both these examples still contain a marked difference in form (and in the former case even an opposition) which can account for the fact that they are able to form good opposite pairs. This is not the case with the antagonyms discussed in 5.2.2. The cross-linguistic evidence shows that the question of whether two opposing concepts are encoded antonymically or antagonymically is not dependent on the type of opposition and it is not the case that the opposition itself is intrinsically 'weaker' in antagonyms than in antonyms since the same concepts may be encoded antonymically in one language and antagonymically in another. However, in the case of an antagonym, the conceptual opposition is encoded as distinct sub-senses of the same lexical item. While antonymy is generally considered a relation between sub-senses of two lexemes rather than between the lexemes themselves (cf. for example K. J. Miller 1998: 49), it seems that the fact that two opposing concepts are encoded in the same lexical form weakens the overt opposition of these two concepts. Speakers do not 'automatically' perceive antagonyms as opposites since the lexical aspect – the encoding of two opposing concepts in two lexical forms which are frequently used together, especially in contrastive constructions (cf. Jones et al. 2007, Murphy 2006) – is missing and the lack of overt difference between the lexical forms makes the difference **too** minimal.

An analogy to phonology serves as a useful illustration, since the phoneme pairs /p/ vs. /o/, /p/ vs. /l/, /p/ vs. /b/ and /p/ vs. /p$^h$/ all differ to certain degrees and some are seen as more strongly opposed than others despite the fact that the first three pairs all differ along the 'voiced:voiceless distinction'. However, the contrast is strongest in /p/ vs. /b/ because these two phonemes are minimally different. The last example is slightly different and while the two phonemes also differ along a salient axis, their 'oppositeness' is weakened by the similarity in form (and by the context the preceding three pairs set) and is thus perceived as less antonymous.

The case of antagonyms which are perceived as only weakly antonymous, if at all, seems to provide support for the claim that antonymy is strongly centred on a lexical level. However, it may simply mean that the conceptual opposition itself loses salience because there is not enough of a distinction between the two lexical items which encode the opposite concepts. It is a binary relation and thus requires **two** opposing items rather than **one** with meanings which are opposed on a more specific level. Minimal difference is a crucial criterion for good opposition which relates strongly to purity of opposition, but it seems equally important that the encodings of the two opposed concepts are sufficiently different.

#### *6.1.1.2 Purity of opposition*

This criterion which was proposed by Cruse (1986: 262) as well as by other scholars of opposition is clearly very influential and it ties in with other criteria discussed below. This section will review the evidence considered in the previous three chapters which relates to this criterion and will also incorporate the discussion of the ease of construction of a unilateral scale (Cruse 1986: 262). Both of these criteria are also discussed by Herrmann et al. (1986: 134f.) who state that clarity of dimensions is a crucial factor in antonym judgements and claim that the clearer the relation is, the better the antonym pairing will be.

The bulk of the evidence for the purity criterion comes from the analysis of the GENDER pairs while the pairs which require a more elaborate thought process for the construction of a binary scale are more varied (e.g. *work:play*, *credit:debit*, *landline:mobile*). Most of the 'impure' pairs in the present study are nominal although verbal pairs are also possible (e.g. *walk:run* or *meander:sprint* based on *slow:fast*) and there may even be adjectival pairs which are less pure than the 'base pair' they are derived from (e.g. *feminine:masculine?*). Purity of opposition seems to be a very reliable criterion since it can be seen how the antonymic strength is reduced by the addition of a larger number of features overall. This is closely related to the complexity hypothesis as well as to word class and will be discussed in greater detail in 6.1.3 below.

The construction of a clear scale along which the two concepts in a word pair differ is a key factor which allows speakers to assess the oppositeness of the features in question. However, the ease with which such a scale can be constructed differs considerably between pairs of words. One factor which makes gradable adjectival pairs such good examples of opposition is the fact that there is no need to construct a scale since this comes 'ready-made' as it is an integral part of the concepts. The different degrees of difficulty in constructing this scale are reflected in the antonymic strength for the pairs in question. The ease with which this scale is construed is also influenced very strongly by context since, if certain features are foregrounded, this may facilitate or inhibit the construal of a certain scale since the context will facilitate the selection of a certain sub-sense of a polysemous lexeme, for example. This factor is also incorporated into the discussion of the holistic cognitive approach to the assessment of antonymic strength proposed in 6.1.3.

#### *6.1.1.3 Symmetry*

Symmetry is, of course, a criterion that only applies to the sub-group of gradable opposites since the existence of a scale is a prerequisite for this feature. Therefore, it will only be considered in the context of the category of gradable antonyms. The importance of the symmetrical distribution of the members of an antonym pair along a scale is clear from all three types of data. Symmetrical pairs co-occur more frequently, are rated higher in the GOE-rating and are recognised faster than their asymmetrical counterparts. Three fields, each centred around a base pair, were constructed to investigate this feature (as illustrated in Figure 3.1) and these three clusters have been discussed repeatedly from different perspectives and in the light of different evidence.

In the GOE-rating data, most symmetrical pairs score better than the asymmetrical ones but it emerges that the symmetrical distribution along the scale in relation to the mid-point is not the only criterion affecting the antonymic strength as far as distribution is concerned. The overall distance from the midpoint also plays a role: there seems to be an ideal distance from the midpoint which coincides with the 'basic level' of the scale (e.g. *hot:cold* on the TEMPERATURE continuum). These are generally the pairs made up from the most common lexemes which have the greatest overall range. The pairs which are here called 'basic level' pairs are also those which are first acquired by children, a factor which is considered an integral part of the basic level distinction (cf. Rosch 1978 and Ungerer & Schmid 2006). It is interesting that the implied neutral temperature (neither hot nor cold), like many neutral points (e.g. on the MERIT scale), is not lexicalised since it would be logical to assume this as the basic level. The neutral case – or midpoint – is, naturally, not fixed since different contexts presuppose different ‘neutral’ settings, for example for temperature or, taking a classical example, size. The neutral size of an elephant is very different from the neutral size of a mug, for example. Furthermore, the GOE-rating results confirm previous claims that the members of the pair are required to be on opposite sides of the midpoint rather than on the same side (Osgood et al. 1957, Herrmann & Chaffin 1986) and ideally equidistant from the mid-point of the dimension (Herrmann et al. 1986: 134f.). Temperature, for example, could also be constructed as two scales, one of heat and one of coldness. Thus concepts which form good opposites cannot be from the same side of a scale and are ideally a similar distance away from the beginning of the scale. This would explain the bad results for pairs like *tepid:lukewarm* since most native speakers would place both these concepts on the HEAT continuum rather than the COLD continuum which means they are on the same side of the mid-point of the temperature scale. Thus, in addition to the symmetrical distribution of the two concepts, overall distance from the mid-point has an influence on antonymic strength.[1]

The data from the German judgement task seems to indicate that symmetry is a more important criterion in German and this phenomenon is supported further by the German behavioural data (cf. 5.4.5.1). The English data, as discussed above, only shows the significance of symmetry as a feature in the subset analysis of gradable adjectives whereas in the German data symmetry displays significance elsewhere as well. However, the question of whether there really are different language-specific weightings of certain criteria would have to be investigated in a separate, specifically designed task.

#### *6.1.1.4 Semantic range*

The criterion of semantic range, which includes Cruse’s match of non-propositional meaning (1986: 262), has been discussed in the analysis of the judgement task as well as in the discussion of the contrastive data in Chapter 5. The greatest support for the importance of this criterion comes from the contrastive analysis and the influence any difference in the overlap of semantic range has on antonymic strength. One example of the analysis of these differences is discussed in 5.2.3.1 when the TEMPERATURE scale is considered contrastively. The lack of a translation equivalent for the pair *boiling:freezing* necessitated certain concessions when

[1] Another factor to consider here is markedness. The highly symmetrical pairs (e.g. *good:bad*) often consist of one marked and one unmarked member which creates an asymmetry. These pairs would have to be investigated in detail to see if markedness plays a role in antonym judgements.

translating this pair into German. The differences in semantic range between *boiling* and *kochend*, as well as the lack of relatedness and symmetry between the two lexemes in the pair *kochend:eisig*, led to a considerably lower GOE-rating score for this pair than for the English counterpart.

Furthermore, all asymmetrical pairs, especially those which contain one member of the base pair, differ in semantic range since they cover a different part of the scale and the range of the base pairs is very often much broader than those of the lexemes on the more extreme parts of the scale. This is, however, also strongly determined by context. In certain contexts, pairs which would be considered decidedly mismatched in terms of semantic range display extremely high antonymic strength. One example for this would be the pair *open:laproscopic* (cf. Paradis et al. 2007) where the first member is an extremely frequent lexeme with broad semantic range while the second member is a very restricted and highly technical lexeme. This pair obtained extremely high co-occurrence scores in the corpus study carried out by Paradis et al. 2007 since it is an excellent opposite pair in a surgical context and appears with extremely high frequency in contrastive constructions.

Semantic range, however, is another criterion which is determined by the internal structure of the two concepts in question. The more closely the structures and features which determine these concepts match, the better the antonymic relation will be – provided, of course, that these features contain a binary opposition. Pairs like *hot:cold* for example are also further strengthened by the opposition of their metaphorical extensions where one member of the pair has an additional, metaphorical meaning (e.g. *hot* 'stolen') and the meaning of the other member of the pair is extended to cover the matching opposing concept (e.g. *They waited for the goods to cool down*).

Pairs which share a larger amount of semantic range will cover the same contexts and be part of the same register, which ensures more opportunity for co-occurrence and thus links this criterion to that of associative strength, discussed in the following section.

#### *6.1.1.5 Associative strength*

This feature incorporates both frequency of co-occurrence and breadth of co-occurrence as well as other associative measures (t-score) used in the present study. This section will summarise the effect the degree of associative strength of an opposite pair has on its antonymic strength and what the relationship between those two measures is. The question of whether the root of the antonymic relationship lies in the degree of associative strength of the lexemes involved in the relation or in the degree of opposition of the concepts involved or, indeed, a combination of the two, will not be discussed in this section but will be considered in 6.2.1.

In the judgement task data introduced in Chapter 3 as well as in the contrastive discussion of the English and German data in Chapter 5, the results clearly show that, while a rating which implies high antonymic strength most often goes hand in hand with a high associative score (both co-occurrence measures from the BNC and EAT elicitation measures), it is not a prerequisite for a good opposite pair to have a high t-score. It is particularly obvious that a high t-score alone does not result in high antonymic strength in the discussions of the converse pairs and the gender cluster. Here, nominal pairs which obtained an extremely high t-score (e.g. *husband:wife, mother:father, king:queen*) score very low indeed in the GOE-rating and are not considered particularly good examples of opposition. The fact that a very high GOE-rating often correlates with a high t-score, while the reverse scenario does not seem to be as reliable, seems to indicate that frequent co-occurrence of lexical opposites is a consequence of the antonymic strength rather than its cause. This is an indication that the methodology used for the judgement task indeed provides a representative measure of antonymic strength.

In the behavioural study, the picture is less clear since the demands of the task are different. The fast presentation of the stimuli and the time constraints on the reaction require a different type of reaction from that in the previous task. The data here shows that associative strength plays a more important role in this type of task than in a self-paced judgement task. If associative strength had as great an influence on antonymic strength as is sometimes attributed to it in the literature (cf. Gross et al. 1989, Miller 1990, Justeson & Katz 1991), there should be a very high level of correspondence between all three measures (t-scores, GOE-rating and meanRT). However, the influence of t-score on the independent variable in the behavioural study is considerably greater than in the judgement task. Had this only applied to very good antonyms, it would most likely not have been noticeable since they would have obtained fast reaction times in any case and the higher associative strength would simply have boosted these times further. However, some of the pairs included in the study were not judged to be excellent opposites but did have an extremely high t-score (cf. examples in the previous paragraph). These pairs obtained a surprisingly high meanRT in the behavioural study. Thus it seems obvious that the degree of influence of associative strength on antonymic judgements differs according to the task in question.

The question of whether the absolute value of the t-score had an effect or whether only very high t-scores and thus very high associative strength influences the results in the behavioural task was raised in the discussion of the data in 4.3.1.2. Often, in the analysis of the sub-groups of pairs those with either **very high** and **high** t-scores or **high** and **medium** t-scores displayed no significant difference in the data, and in the overall analysis of the data those pairs with **medium** and **low** t-scores did not differ significantly from each other. It thus seems as if there is a co-occurrence threshold which is needed to provide the level of facilitation seen in the nominal examples discussed above. Thus if there are pairs included in the dataset which have a very high t-score but do not have the same degree of antonymic strength as conventional good opposites, the results of a behavioural task will be skewed by the disproportionately fast reaction times obtained by these pairs.

K.J. Miller's claim that '[the mutuality of association] seems to be acquired as a consequence of these pairs of words being used together in the same phrases and sentences' (1998: 48) seems to be looking at the issues from the wrong perspective, since the data shows the high degree of co-occurrence of the members of an opposite pair to be a result of their close association rather than a determining factor of it. However, pairs like *mother:father* display the same degree of co-occurrence as many excellent opposites (as, in fact, do many collocations) but this in itself does not make their relation antonymic. Two factors are at work here and it is important that these be kept separate.

Thus it can be concluded that associative strength does play an important role in the behavioural data as it seems to lead to quicker activation and recognition of items which are very strongly associatively related. Despite the fact that their antonymic strength (as judged by GOE-rating) is not as strong as that of other pairs, they result in similar reaction times. There seems to be a component of automatic activation which is facilitated by the close associative relation and lays the groundwork for faster recognition and decision in tasks of this type (cf. 7.2).

### 6.1.2 Additional 'old' criteria

#### *6.1.2.1 Inherent binarity*

This section will address a criterion which has not been discussed at length in the present work since it is difficult to quantify and its application seems rather subjective. Inherent binarity was put forward by Cruse (1986) as a criterion for good oppositions after discussing the distinction between *black:white* and the other colour terms and the antonym pair *happy:sad* as opposed to other adjectives describing emotion (e.g. *angry*, *calm*). He comes to the following conclusion about the criterion of inherent binarity:

> An inherently binary contrast, in other words, would be whose binarity was logically necessary. […] Many relations based on the idea of a uni-dimensional scale or axis are inherently binary, and it may be that all inherently binary contrasts are describable in this way. (Cruse 1986: 259)

This criterion seems to tie in with scalarity of certain opposite pairs which, as discussed above (cf. 6.1.1), is one feature out of Lyons' and Cruse's original criteria and with definitions of antonymy which seems to affect speaker judgement. If an opposite pair is distributed evenly across the mid-point of a certain scale it seems to be much more easily identified as an opposite pair than when the property it encodes is non-scalar. This does not hold for all cases since there are a number of strongly conventionalised pairings which are not gradable in any way but are nevertheless judged to be good opposites.

Furthermore, the criterion of inherent binarity provides an explanation for the lack of antonymic strength between co-hyponyms such as *cat:dog* or *red:green* which are weakly opposed and can be used as full antonym pairs in certain contexts but would not usually be considered antonyms 'proper' (i.e. canonical antonyms) by native speakers of English. According to Cruse (1986: 259) they lack inherent binarity, and this weakens their antonymic strength.

The question of whether Cruse intended inherent binarity as a lexical or a semantic criterion remains unanswered. However, it seems logical to assume inherent binarity to be a criterion which works at the level of conceptual processing rather than lexical processing since both pairs given as examples above (*cat:dog* and *red:green*) are extremely frequently associated in text (cf. Appendix 3 for co-occurrence data) but this strong associative relation nevertheless does not result in a strong antonymic link between the two concepts. Thus, this does not appear to be a relation guided by the association between two lexical items (contrary to claims by, for example, the WordNet research group).

It can be considered a prerequisite for canonical antonyms to be inherently binary, and the stronger this binarity is, the more it strengthens the antonymicity (cf. Cruse 1986: 260) of any given opposite pair. However, in the category of non-canonical opposition, inherent binarity is less important. It is a concept that is difficult to quantify and measure empirically, which was the main reason for not including it in the list of criteria to be investigated. Furthermore, the fact that opposites that are inherently binary are better than those which are not (especially out of a particular context) should not come as a surprise since, if this was not the case, any two co-hyponyms would be considered potential candidates for lexical opposition.

*6.1.2.2 Morphological relatedness*

The evidence gathered from the GOE-rating tasks in English and German as well as the behavioural experiments is re-evaluated here in terms of the theoretical implications for the importance of morphological relatedness as a criterion for good opposition. The first question which has to be asked is what kinds of pairs should be considered ‘morphologically related’ in this context. The questionnaires and the behavioural tasks all contain pairs which display more than one type of morphological relatedness:

a) prefixation with a negative prefix (e.g *happy-**un**happy*)
b) suffixation with a gendered/relational suffix (e.g. act**or**:act**ress** or employ**er**:employ**ee**)
c) affixation with contrasting or opposite affixes (e.g. ***in**clude:**ex**clude* or *help**ful**:help**less***)
d) pseudo-antonyms (*flammable:**in**flammable* and *easy:**un**easy*)

Of the 23 morphologically related pairs of type (a), only two are not in the group of excellent antonymy (1.0 – 1.8) in the English GOE-rating: *married:unmarried* (1.965) and *interested:disinterested* (2.05). There is no morphologically related pair in the top ten results and only one (*include:exclude* – 1.15) in the top fifteen. This shows that, in the judgement task at least, morphologically related adjectives form a more homogeneous category than their unrelated counterparts. Muehleisen (1997) states that related adjectives show more overlap of semantic range than those which are morphologically unrelated. One could assume that this is the case because the related adjective is automatically assumed to cover the same range as the base adjective since it simply provides the negation of the base concept. This would mean that not only the similarity in form but also the larger overlap in meaning and usage which is tied to a largely overlapping semantic range both contribute significantly to the greater antonymic strength and more consistent nature of the results observed in these pairs. As already noted in 3.1.3, these related pairs can be seen almost as a positive control – an example of the rating an antonym pair should receive in a judgement task. The influence of relatedness on judgements can be seen in the results of the two pseudo-antonymic pairs: *easy:uneasy* (3.675) and *flammable:inflammable* (3.455). While these results are by no means ‘good’, they are significantly better than expected since the former is usually considered barely antonymous (other than in a very specific context) and the members of the latter pair are technically synonyms. These pairs display a larger standard deviation (participants chose either the very top or bottom of the scale with few responses at 3 and 4) which indicates that some participants were misled by the morphological relationship between the two members of a pair whereas others were not.

The behavioural data taken from Experiment 1 does not reveal any clear patterns in a comparison of the morphologically related pairs to their unrelated counterparts. If anything, the related pairs display a marginally longer RT which could be accounted for by the increased item length. However, while item length may have been a contributing factor, it is clear from the results that it cannot have been the only, and by no means the most influential criterion, since some of the shorter pairs (e.g. *odd:even*) display a considerably longer RT than the corresponding related pair (*even:uneven*). Lyons's claim (1977: 275) discussed in 6.1.1.1 that unrelated pairs are better antonyms because they have less in common may not be entirely accurate since related antonym pairs display great antonymic strength and the measures in the present work do not allow for a clear separation of the influence of this particular criterion, especially since the factor of morphological relatedness may have been influenced by repetition priming in Experiment 2.

To further investigate the effect of morphological relatedness on antonym canonicity, it would be interesting to design an elicitation experiment with those lexemes as stimuli which can take either a morphologically related or an unrelated partner to see whether the results display a bias in either direction or whether both the related and unrelated forms are elicited with equal frequency. In the GOE-rating data, when the top two categories (excellent and good) are analysed, the ratings for morphologically unrelated and related pairs do not differ significantly and the average GOE-score is almost identical (1.68 (related) vs. 1.64 (unrelated)). This analysis is only intended to be illustrative since the number of unrelated items was vastly larger in the questionnaire and the results are thus bound to be skewed.

*6.1.2.3 Sequencing*

The analyses of the data from both the GOE-rating and the behavioural tasks have shown that the order in which the members of an antonym pair are presented does not have an influence overall. Despite this finding, sequencing of opposite pairs has been investigated by several scholars and criteria have been proposed which explain the preference in sequence of certain pairs (cf. Cruse & Togia 1995, Jones 2002 and Croft & Cruse 2004). There is very often a sense that the more salient of the two properties, which is usually the one which is unmarked (cf. Jones 2002), is preferred as the first member of the pair (e.g. *long:short*, *hot:cold*). In this section, individual pairs are considered in order to investigate whether the results support these criteria despite the fact that no overall significance of the sequence criterion can be established in either dataset in English or German.

The GOE-rating data has not been treated with the same statistical methods as the behavioural data; however, the table below lists 20 pairs (out of 210) which display a discrepancy in their GOE-scores between the preferred and dispreferred sequence. These pairs show a clear preference for one sequence in the data. The table also includes the raw co-occurrence frequencies by sequence to illustrate that the GOE-rating data does not necessarily match the co-occurrence data. Only six (marked with * in Table 6.1) of the 20 pairs show the bias in favour of the sequence which is predicted by the GOE-rating results. This, however, does not mean that either of the measures are unreliable but, when the table is considered in greater detail, it is clear that there are only very few cases where the GOE-rating discrepancy goes in the opposite direction to the one indicated by the co-occurrence data. In one case, *damage:repair*, it is surprising that the GOE-rating matches the co-occurrence analysis since the higher frequency may be due to mis-tagging of *damage* (noun) as *damage* (verb) and cases of *to repair damage* may have been counted in the frequency count for this pair.

**Table 6.1** *Pairs with discrepancies greater than .5 in the GOE-rating*

| Word 1 | Word 2 | Freq (1-2) | Freq (2-1) | GOE (1-2) | GOE (2-1) |
|---|---|---|---|---|---|
| ask | answer | 30 | 25 | 2.9 | 2.55 |
| buy | spend | 5 | 5 | 5.15 | 6.05 |
| buy | sell | 157 | 48 | 1.93 | 1.43 |
| child | parent* | 66 | 188 | 4.05 | 3.1 |
| chilly | steaming | 0 | 0 | 3.08 | 4.05 |
| correct | mistaken | 0 | 0 | 2.73 | 1.95 |
| current | former | 21 | 17 | 3.43 | 2.95 |
| damage | repair* | 11 | 89 | 2.75 | 1.98 |
| even | uneven | 0 | 0 | 1.28 | 1.63 |
| figuratively | literally | 2 | 3 | 2.23 | 1.7 |
| finish | continue | 0 | 3 | 4.23 | 4.75 |
| good | disobedient | 0 | 0 | 3.05 | 3.7 |
| inaccurate | accurate | 0 | 5 | 1.28 | 1.08 |
| lend | borrow* | 7 | 15 | 2.9 | 1.98 |
| male | female* | 903 | 181 | 1.88 | 2.18 |
| present | future* | 179 | 27 | 3.48 | 4.2 |
| present | past | 107 | 494 | 2.43 | 3.13 |
| sour | sweet | 2 | 40 | 2.33 | 3.03 |
| town | country* | 429 | 82 | 3.18 | 4.25 |
| undress | dress | 0 | 5 | 1.68 | 1.35 |

When the salience criterion (Croft & Cruse 2004: 19) is applied to the pairs above, it is immediately obvious how some of the results come about: in both *present:future* and *present:past*, *present* is the preferred first item and, as it is the starting point for our thinking, the more salient concept. However, *current:former* does not correspond to this format. All morphologically related pairs in the table are in the order which is expected with the derived form appearing as the second item of the pair in the preferred sequence.

In the behavioural data, there are eleven pairs which obtained results differing in meanRT by more than 100ms. These pairs and their two mean reaction times are illustrated in the chart in Figure 6.1. The pairs are shown in their preferred order (as determined by meanRT) which means that the darker line shows the results for the pairs as they are displayed while the lighter line gives the times for the pairs in reverse order. Apart from two pairs, *defend:attack* and *cool:warm*, all pairs are in the order which would be expected according to the criteria presented in the literature (cf. also 2.2.2) – *more* before *less*, *big* before *small* and *male* before *female*.

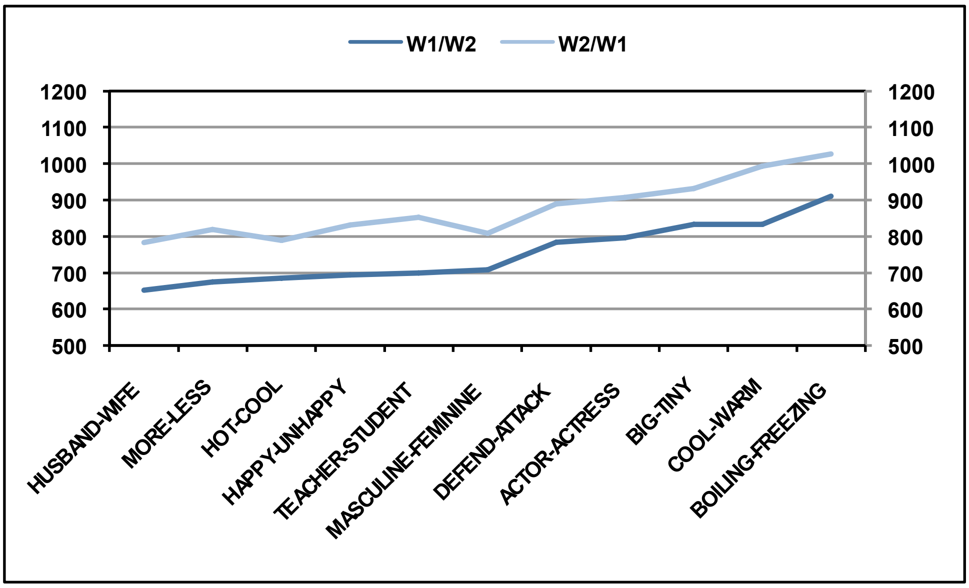



**Figure 6.1** *MeanRT (in ms) for pairs which differ by more than 100ms (Exp. 1)*

Two final observations can be made when considering the chart in Figure 6.1: firstly, the distance between the two graphs is relatively consistent with the dispreferred pair obtaining reaction times which are slower by a similar amount of time. Secondly, it is striking that none of the pairs in either of the lists (Table 6.1 and Figure 6.1) is one of the highly canonical pairs despite the fact that many of these pairs seem to have a preference in sequence when they are being named and many native speakers find it relatively easy to decide which sequence they would prefer in some cases and rather more challenging in others. The results show that, despite this latent preference, the sequence is not of great overall importance in antonym judgements and the antonymic strength of very good antonym pairs seems to be affected even less by the order in which the pair is presented.

### 6.1.3 A holistic approach – antonymy as a cognitive construct

The new criteria presented throughout the analyses in the previous chapters have all emerged from the data-driven approach to antonymy. Some of these criteria are not new per se but have a different focus than before and are moulded and supported by the empirical evidence collected for this study. Rather than breaking the factors down into independent but interrelated criteria, they are presented as being based on one central tenet. Many of the features presented in the previous sections are incorporated into a holistic approach which attempts to account for the degrees of canonicity observed in the collated data, as well as to inform our understanding of how lexical opposition is construed. All criteria presented here are strongly related and assume antonymy to be a prototype category in the cognitive sense with an internal structure closely matching that of other such categories and is, like those, strongly influenced by context.

This section considers the evidence for the hypothesis introduced in 3.4.3 which states that the antonymic strength of opposite concepts or categories is dependent on a certain degree of overlap and similarity between the two concepts involved. Word class, the complexity hypothesis and the question of salience are discussed below and an integration of the other factors which influence antonymic strength into this model is attempted.

The basic theoretical principles which underpin the proposed approach to antonym canonicity are firmly anchored in the cognitive framework of semantic analysis and meaning construction (cf. Rosch et al. 1976, Lakoff 1987, Langacker 1987, Geeraerts 1994, Evans & Green 2006). The constructs which are used here are those of category structure, prototype effects, salience and foregrounding and idealised cognitive models (ICMs) which underlie concepts and link them in unbounded networks. The internal structure of these categories, which represent the concepts encoded in the members of an antonym pair, is of prime importance. The flexibility of category-internal structure is well-attested in psychological and now psycholinguistic research (cf. Lakoff 1987: 83ff.) and this flexibility is crucial here.

Each member of an antonym pair is a conceptual category which is determined by certain attributes and whose structure is influenced by underlying cognitive models which may be cultural, contextual or dependent on register. The match between the structure of the two concepts in an opposite pair determines its antonymic strength. Within this framework, there are several factors which influence the quality of this match and thus the degree to which concepts are perceived as antonymic.

*6.1.3.1 Complexity of category structure*

In Chapter 3, following the analysis of the judgement task data, the complexity hypothesis is introduced alongside the claim that lexemes which encode more complex categories are less likely to be judged to be good antonyms. This ties in directly with purity of opposition and also with the criterion of word class introduced in Chapter 3 (cf. 3.4.2). Certain lexical categories (e.g. adjectives) are, by definition, less likely to have a complex category structure linked to an underlying cognitive model since their structure is simpler and more stable. In terms of antonymic strength, this is an advantage since, as in purity of opposition above (cf. 6.1.1.2), the fewer additional attributes cloud the antonymic features, the stronger the opposition will be. However, it is not simply the complexity of the internal structure of the categories involved but also the degree to which the structures of these categories match – both in the attributes which define the structure as well as in the underlying models – which determine much of the meaning construction of the individual categories (cf. Lakoff 1987: 83ff). Thus, categories which share a large number of attributes overall and are of a similar density in structure (criterion of **minimal difference**) will make better antonym pairs than those which differ more considerably in those two aspects. Similarity in the underlying models is also an important criterion (cf. **semantic range** and **match of non-propositional meaning**) as the underlying models determine the contexts the lexeme is likely to appear in; this, in turn, will influence the rate of co-occurrence of the two items.

*6.1.3.2 Word class*

Adjectives, which make the best antonym pairs overall (cf. 2.2.2 for example), show the least amount of complexity among all lexical categories investigated in this study. There were not enough prepositional pairs to determine their status on the complexity continuum but they should be of a similar or identical status to adjectives, having a relatively simple category structure.

Nouns and verbs, neither of which co-occur as much in the same syntactic environments as adjectives do (cf. 2.5.3.1 Fellbaum 1995) generally have a more complex category structure and are also linked to underlying cognitive models of greater density. According to the complexity hypothesis introduced in 3.4.3, the larger number of overall features in nominal and verbal pairs detracts from the antonymic features which determine the opposition and thus reduces its salience. Thus, the more complex the category and the greater the number of determining features is, the weaker the antonymic strength of the pair in question. It is evident that this approach incorporates the 'old' criterion of purity of opposition since *pure* can here be seen as a synonym of *simple*. As previously stated, complexity is not necessarily tied to word class but the pairs under investigation in this study show the pattern discussed above.

#### *6.1.3.3 Construal of a scale*

The construal of a scale along which the opposition is set up would then be easiest in simpler categories since the opposed concepts are less overshadowed by other features and will thus be immediately salient. Naturally, there are pairs for which the scale is more easily identifiable – as mentioned above, all gradable opposites come with a 'ready-made' scale which is pre-construed and highly conventionalised, while others require more effort for this construal regardless of the complexity of the category structure. Finally, there are also words for which such a construal is impossible and which thus do not form an opposite. The less cluttered the internal structure of the two categories involved, the easier it will be to identify the scale along which they differ and, in this approach, the easier it will be for speakers to make antonym judgements about the pair of categories in question.

#### *6.1.3.4 Salience*

The last, and most important, determining factor in this model is that of salience (cf. for example Rosch 1975, Rosch & Mervis 1975). The attributes or features of the concepts in the antonymic relation must differ along a salient dimension for the opposition to be recognised. Non-salient opposition is often construed as a comic mechanism when one minor feature which weakly contrasts with another is chosen to create an artificial opposite pair. However, this is not the type of opposition which is investigated here. The degree of salience of certain features is also directly linked to the complexity of the category structure since, if there are fewer choices, it is easier to establish salience in a less complex environment. If the categories in question are rich in attributes, the dimension along which the two pairs differ must be of greater salience than in a category with a simpler structure and thus the opposition may be more difficult to establish. This was seen in many of the nominal and verbal pairs above which overall show greater complexity in their category structure and which are also determined by the underlying models to a greater degree than their adjectival counterparts. The reason why there are some nominal and verbal opposite pairs, for example, which were judged to be excellent opposites despite the fact that pairs from these lexical categories are generally found lower on the canonicity continuum is that the opposed properties in these concepts were the most salient attributes in the neutral setting they were presented in (e.g. *giant:dwarf* and *Tod:Leben*).

However, antonymic salience is not determined only by the (inherent) binarity of two features of a concept which may be more or less easily seen depending on the complexity of the overall category structure. In addition, contextual, cultural and register effects all have considerable influence on the salience of a certain property. One advantage of this holistic, cognitive approach to the factors which determine antonymic strength is that this model can account for the context phenomena found in antonym pairs. The context determines which cognitive models underpin a concept and thus which of its features are of particular importance in any given use of the concept. If the antonymic features are foregrounded by a particular context, the pair is likely to be considered to have more antonymic strength in this particular context since the antonymic features are those which are of the greatest salience in that particular usage. The poem by Wilbur at the beginning of Chapter 1 is a good example since it illustrates this context dependence by first introducing the opposite of *white* in a neutral context – *black* – but then modifying the context to that of EGGS in which case the opposite of *white* changes to *yolk*. Without the specific context provided and the encyclopaedic knowledge provided by the models which underlie and link these concepts, it would require much greater effort to construct the binary opposition between *white* and *yolk* and they may only be judged very faintly antonymic.

These context-dependent opposites are, of course, not part of the cluster of canonical opposition since they are conceptually opposed but do not fulfil the added surface criteria which link particularly strong opposites in addition to their cognitive opposition: morphological relatedness and associative strength. However, both of these criteria are tied to the cognitive relationship between the conceptual opposites and thus also determined by it. A morphologically related opposite pair will match very closely in category structure and differ only along a salient dimension since these pairs are, in a sense, constructed to be opposites. This also works on an ad-hoc basis by the modification of a lexeme with a reversative prefix which turns the concept encoded by this lexeme into its reverse. A high degree of associative strength is dependent on an excellent match of the concepts since this allows for usage in the same contexts. Register is also accounted for in a similar manner since the cognitive models which underpin the concepts as well as the conceptual structure will change according to register, resulting in a change of the level of salience of some of the features involved.

The mechanism would have to be such that the concepts are compared, a common dimension is found (construction of a uni-lateral scale) and then the features which are part of the cognitive model of each member of the pair are analysed and matched to see whether this results in a binary contrast. It seems intuitively right that this would be an easier process and thus more swiftly accomplished if the attributes in question were foregrounded either because of the structure of the category, i.e. their centrality of position in the conceptual structure, or by the context which the concept is tied into.

#### *6.1.3.5 Ranking of criteria*

After the division of the criteria into cognitive criteria (which influence the conceptual opposition between the meanings of the two lexemes which form an antonym pair) and lexical criteria (morphological relatedness and associative strength), a ranking of the criteria proposed in 2.5 seems ill-advised. The cognitive criteria all form part of the same model and good opposition can only be established through their interplay but no firm ranking can be established. The most important criterion is salience, but this is determined by many of the other factors introduced and investigated above and thus does not form a stand-alone factor. The two lexical criteria are not a necessity but they do influence antonym judgements by native speakers. They are, however, related to a different level of the opposition. Both will be taken into account in the discussion of the nature of antonymy as a lexical or semantic relation in 6.2.2 and 6.2.3 below.

## 6.2 Antonymy in the mental lexicon

The empirical data gathered for the present research gives rise to certain hypotheses about the nature of the relation of antonymy in the mental lexicon. The questions which were raised in the earlier stages of this study and which will now be re-evaluated are those of the influence of antonym type on antonym judgements, the question of whether antonymy is construed 'online' as required or whether it is a relation which is stored and whose strength is determined by the entrenchment of the relation of the two lexemes which form an antonym pair. The third and last question which will be considered in this section is whether antonymy is a lexical or a semantic/cognitive relation and what evidence the data contributes to this discussion. Previous standpoints on the various issues under discussion which were introduced in Chapter 2 are re-evaluated below in light of the evidence provided by the present study.

### 6.2.1 Psychological reality of antonym types

The categorisations of different types of antonyms which were proposed in Lyons 1977 and Cruse 1986 have provided a starting point not just for the present research but also for several other studies which all take note of the logical differences which give rise to these categories and use them to structure their research. While the distinctions between the categories and subcategories proposed in the two classifications above are certainly appropriate from a logical and theoretical (and even a syntactic) perspective, the question has been raised by several scholars whether these differences affect speakers' judgements of antonymic strength.

Jones (2002; cf. 2.2.2) has investigated whether gradability, which constitutes one of the key factors in the categorisations above, affects the textual function of antonyms. He concludes that gradability does not have an obvious effect on usage and that the similarities between gradable and non-gradable opposites are more pertinent than the differences. Gradability did not emerge as an important criterion from any of the data gathered despite the fact that their study did include more non-gradable than gradable pairs. This was accounted for in all analyses and the difference between the GOE-rating and behavioural results for gradable and non-gradable opposites is not significant. However, some of the best pairs in the study, for example *good:bad*, *hot:cold* and *large:small*, are gradable. Nevertheless, only these very central members of the category of opposition score very high since the factor of symmetry comes into play here which makes gradable, non-symmetrical pairs weaker in terms of antonymic strength than non-gradable opposites. These results tie in with those of the corpus study conducted by Jones (2002) which states that gradability does not affect the way in which opposites are used in text (or speech; cf. Jones 2006).

In 2.3.1, where cognitive models of antonymy are considered, many different systems of scalar opposites are proposed within the category of gradable opposition. However, it is not clear that the difference in type of the underlying scale (cf. 2.3.1; Cruse & Togia 1995) has an influence on the ease of construal which, in turn, influences antonymic strength. The question of whether these differences in construal influence speakers' judgements is raised by Croft & Cruse (2004). In the data collected in the present study there does not seem to be a difference in the results for pairs which belong to the different systems of constructions introduced by Cruse & Togia 1995. All central pairs in the study which correspond to any of the scales in their research show the same degree of very high antonym canonicity while the pairs which display a lesser degree of strength do so for other reasons.

The GOE-rating and behavioural results in the present study both showed differences between some of the antonym types. In the judgement task, the only difference was that between converses on the one hand, and antonyms and complementaries on the other. This applies to English and German equally, while in the behavioural task there is a difference between the results in the two languages. The English results display the same pattern as the judgement task data with the only significant difference being seen between converses and the two other antonym types, while the German results show significant differences between all three types. As discussed in 5.4.3, the differences between antonyms and converses are smaller but nevertheless significant. However, since the German dataset was substantially smaller than the English one, it is possible that these effects are exaggerated by particular word pairs which were less well matched than in the English version of the experiment (cf. 5.1.3).

The results presented above clearly indicate that there is no discernible difference between gradable and non-gradable antonyms in terms of antonymic strength and that the theoretical difference in type of antonymic relation does not translate into a real difference in usage. However, in the case of converses, the evidence points in a different direction. Converses are significantly different in the GOE-rating results in both languages as well as in the behavioural study. This observation cannot be accounted for by differences in co-occurrence since, in an ANOVA of t-score and antonym type, the only significant difference could be seen between antonyms (average t-score of 10.545) on the one hand and converses and complementaries (average t-score 12.705 and 12.277 respectively) on the other. Thus the difference must have a conceptual explanation. The complexity hypothesis can also be extended to account for converses. It is assumed that the more complex logical relation which holds between the two members of a converse pair (cf. 3.4.2) does have an effect on the construal of the scale and thus on the antonymic strength on the pair. According to the discussion in 6.1.3 above, the fact that converses are nominal and verbal rather than adjectival will also influence the degree of antonymic strength of these pairs since the internal structure of the concepts will be more complex.

It can thus be concluded that the only one of the theoretical differences which can be substantiated by experimental data is that between converses and other opposites with a less complex connection. Whether it is indeed the added complexity of the relation or the fact that converses are of a different lexical category cannot be determined from the data collected in this study and would have to be investigated separately in a study which contained nominal and verbal pairs from different antonym types only.

### 6.2.2 Feature matching vs. entrenchment of association

This section returns to Aitchison's question of whether humans 'work things out or look things up' (1994: 95) which was introduced in 4.1.3. The question whether antonymy comes about through feature matching or through a strong associative relation between the two members of a pair which means they are so closely related that the relation does not need to be construed but can simply be looked up has been one of the key issues in this study. Two separate questions will be considered in this section: whether a feature matching approach seems appropriate in the light of the data collected in this study, and what role, if any, the associative relation plays in antonym judgements. Evidence such as that of contextualisation of opposites gathered in other studies (cf. Murphy & Andrew 1993) will be discussed in relation to the present data as well as the methodological question of whether the behavioural data shows the results of antonymic strength or whether the associative strength of each pair will need to be taken into account to achieve an accurate interpretation of the data.

If antonymy is a case of ‘working out’ rather than ‘looking up’, the context-dependence of antonyms, which has been described as highly idiosyncratic (cf. Murphy 2003) can be explained using the model presented in 6.1.3. Feature matching allows for the foregrounding of a particularly salient property in a certain context. An approach based on associative strength would find it problematic to account for the fact that the same opposite pair can be said to have more or less antonymic strength depending on the context it is encountered in since its overall co-occurrence and thus entrenchment would not change and thus would not allow for this flexibility in antonymic strength.

The behavioural experiments described in Chapter 4 are based on the mechanisms of spreading activation and priming which are strongly dependent on associative relations. In a review of a number of experimental papers on priming, Hutchison (2003) investigates the factors which influence semantic priming. He concludes that both feature overlap and associative strength have an influence on semantic priming but that in the case of both synonyms and antonyms an approach based on feature overlap is favoured. This is further supported by recent neurolinguistic studies comparing the processing of synonyms and antonyms (cf. 4.1.3 – Jeon et al. 2009).

The results of the behavioural data discussed in Chapter 4 show that associative strength does play a role since items with very high associative strength seem to elicit faster reaction times than would have been expected from their judgement task results (e.g. *mother:father*, *husband:wife*). These are the cases in which the advantage of high associative strength is obvious as the discrepancy between their antonymic strength and their associative strength – at least in a neutral context – is considerable. It can be assumed that many other pairs which co-occur very frequently will also experience the same advantage but this is less evident since antonymic strength and associative strength are, in many cases, very well matched as there is a strong correlation between the two factors. One hypothesis is that the associative relation is automatically activated and aids the processing of the two items, since the feature matching can then be conducted more efficiently as both members of the pair are already activated and a relationship between the two is already in existence. This hypothesis received support from the results of Experiment 2 which was constructed as a priming experiment, presenting the two members of a pair with a 450ms delay. The reaction times for opposite pairs were significantly shorter in the majority of cases and in some targets the priming paradigm also generated inhibitory priming when the associative relation activated the primary partner and an alternative second lexeme was presented (e.g. *big:little* – cf. 4.4.2.2).

The above discussion provides support for the hypothesis that antonym pairs are constructed rather than stored but that there is a significant influence of associative strength in priming experiments which speeds up the results of those pairs which co-occur particularly frequently and whose relation is conventionalised. This, however, is not directly related to antonymic strength since this process also applies to pairs with high associative strength but low antonymic strength. The difference between conventional 'direct' antonym pairs (e.g. *wet:dry*) and those termed 'indirect' antonyms (e.g. *soggy:parched*) found by Gross et al. (1989) and Gross & Miller (1990), and to a certain extent by Charles et al. (1994), can also be explained by the mechanisms described above since the 'direct' canonical antonyms in their studies were those which would have benefited from the additional associative strength whereas the constructed 'indirect' opposites were less well entrenched and would have had to rely on feature matching only: this would explain the discrepancy in reaction times between the two groups. This is also supported by the neurological data gathered by Gazzaniga & Miller 1989 who find that their data does not support the view of mediated antonymy. Data from an ERP study to investigate the differences in processing between antonym pairs with high associative strength and those whose connection is less well entrenched may be able to provide vital data to further substantiate the hypotheses presented above.

It seems that the results of the behavioural task are influenced strongly by associative strength but that they nevertheless present a good reflection of antonymic strength since there are also a large number of pairs with relatively low t-scores which obtain very fast reaction times. Thus, this task shows that the lexical aspect seems to have an influence on the processing of antonyms and is especially important in canonical cases. What this means with regard to the question of what type of relation antonymy is will be discussed in the following section.

### 6.2.3 Lexical vs. semantic relation

Before taking a look at the internal structure of the category of antonym in the following section and comparing the results of the data collected in this study to the multi-method approach by Paradis et al. 2009, the question of whether antonymy is a lexical or a semantic/conceptual relation is addressed. The preceding sections have stressed the cognitive element in the construal and processing of antonymy and have made a case for the conceptual basis of the relation. However, the influence of lexical aspects on the relatedness of the two items forming an antonym pair has also been noted, especially in relation to the results of the behavioural study. This section will attempt to distinguish between the two aspects and determine the influence of lexical factors on overall canonicity.

To begin with, the distinction between lexical and semantic relation which was introduced in Chapter 2 (cf. 2.4.1) must be clarified. Miller (1990) claimed that antonymy was a relation between word forms rather than between the meanings of those words and the concepts encoded by them. This is what is termed a 'lexical' relation in the literature on antonymy. On the other hand, identification of antonymy as a semantic relation is supported by those scholars who believe that this relation is based on an opposition in conceptual structure and meaning rather than on the pure contrast resulting from the frequent co-occurrence of certain word forms in contrastive contexts. There is evidence in the data for both approaches, and with a more restricted set of stimuli and less diverse empirical methods, the conclusions drawn from the data may well have been (or seemed to be) more easily interpretable in the conventional way (cf. also 7.2).

The conceptual side of the argument receives strong amount of support from the intralingual GOE-rating data which shows considerable discrepancies between the measures of associative strength and the results of the judgement task in a large number of cases (cf. 3.4). However, the most striking evidence in support of the conceptual hypothesis comes from the comparison of the German and English judgement tasks. The results discussed in 5.1 show that most of the 210 opposite pairs in the two languages do not differ by more than 0.5 on the seven-point scale the participants were given. The same pairs are judged to have very high antonymic strength in German and in English despite large differences in the comparison of t-score rankings between the two languages. The great similarity of antonymic strength between the two languages indicates at least a conceptual basis for the relation of antonymy since it seems unlikely that the same concepts would form contrastive relationships of near-identical strength without some cognitive or conceptual trigger. The judgement task data thus provides evidence for antonymy as a semantic relation and shows clear differences between antonymic and associative strength with associative strength being understood as a result of the conceptual opposition rather than as a cause of it. The question why a particular pair of lexemes is chosen to represent the opposed concepts when there are other near-synonymous lexemes available is partly answered above by criteria such as semantic range and a close match of the conceptual structure encoded by a particular lexeme as well as the cognitive models which underpin these concepts.

The behavioural data shows a slightly different picture since the associative aspect seems to have much greater influence here than in the judgement task. It is apparent, however, that there are two different mechanisms at work since great associative strength also affects lexemes with weaker antonymic strength which indicates that, although connected in some way, these are two different factors which both influence the processing and recognition speed of antonym pairs. While this data gives more weight to the associative criterion than that obtained from the judgement task, this does not mean the data is at odds with the idea of antonymy as a conceptual relation. The influence of associative strength is greatest at the top end of the canonicity continuum, especially in the traditional canonical opposites. Thus, the claim made by Miller that 'antonymy is a lexical relation between word forms, not a semantic relation between word meanings' (1990: 242) could be justified if one only took the central category of gradable, adjectival opposites into account, since their 'status' is set apart from the rest of the members of the category of antonymy by the added advantage very high associative strength provides in terms of entrenchment and automatic activation.

However, when the whole spectrum of antonyms along the canonicity continuum is considered, it becomes clear that the conceptual aspect is crucial and is most likely the aspect which determines canonicity and which leads to the frequent co-occurrence and thus high associative strength displayed by the more canonical of the antonym pairs. It can be seen from the discussion above that the different strands of data have contributed different parts of the puzzle which has allowed this study to provide solid evidence for the theory of antonymy as a semantic and conceptual relation. Therefore, it seems of great importance to consider several methodological approaches since the methods which are used will inevitably guide the interpretation of the data (cf. 7.2 for further discussion of the methodological aspect).

## 6.3 Antonymy as a gradient phenomenon – a comparison of evidence

As can be seen from the evidence discussed in the previous sections, the experimental evidence does not unequivocally support a distinction between direct and indirect or canonical and peripheral opposites. A large amount of space and attention has been given to the concept of canonicity in the present work and it seems that the question whether canonicity is a gradable property resulting in an antonym continuum or whether antonym pairs are merely either canonical or not has to be addressed separately from the discussion of the nature of the phenomenon of antonymy and the factors which determine the place of an antonym pair on the canonicity continuum.

The idea of antonymy as a gradient phenomenon is not altogether new since, as already mentioned in 2.2.1, Mettinger (1994: 162) pointed out that encyclopaedic knowledge would increase on a scale depending on antonym types:

> A treatment of opposites from a cognitive point of view would probably have to assume a scale of increasing amount of encyclopaedic knowledge that is necessary for establishing conceptual integrators and differentiators, starting from a zero-value (characteristic of systemic opposites), becoming more with terminological opposites, and reaching a maximum with encyclopaedic opposites.

As seen in 6.1.3.2, this increase of encyclopaedic knowledge seems to be detrimental to antonymic strength so that the necessity of greater encyclopaedic knowledge, which goes hand in hand with increased complexity of category structure, would lead to a decrease in antonymic strength. The three categories proposed by Mettinger (1994) – systemic, terminological and encyclopaedic opposites – could be extended to four by adding contextual opposites as a category which requires even more language-external knowledge than the category of encyclopaedic opposites since the situational and cultural context will also have to be taken into account. This 4-part distinction seems to be extremely representative of the phenomena observed in the data if the categories are not taken as absolute but as relative to the preceding pair, resulting in a continuum from no encyclopaedic knowledge to a large amount of encyclopaedic, cultural and contextual knowledge. This scale would then be directly proportional to an inverse scale of antonymic strength.

Furthermore, Herrmann et al. (1986) conducted judgement tests which led to results that showed clearly that 'goodness of antonymy' appears to be a gradable rather than an absolute phenomenon. If antonymy were divisible into two distinct categories, direct and indirect opposition, possibly even with one mediated by the other (cf. WordNet; Gross & Miller 1990), this should be visible in the experimental results where one should be able to determine a clear cut-off point between canonical and non-canonical opposites in whichever independent variable the experiment relies on. Such a 'drop' in goodness of antonymy has so far not been attested.

One investigation in particular, conducted by Paradis, Willners & Jones 2009, theoretically and empirically resembles the present research and will thus be discussed in detail below since useful insights can be gained from a comparison of the two studies. Paradis et al. 2009 also employed a multi-method approach to the investigation of antonymy but used different techniques. As in the present research, the use of a multi-method approach was considered to be overdue and it was clear that the comparison of the results of different methods would provide valuable starting points for the theoretical and practical investigation of lexical opposition. In the next section, I will introduce the methods and hypotheses proposed by Paradis et al. 2009 before comparing their results to those of the present research.

### 6.3.1 The cognitive prototype model of antonymy (Paradis et al. 2009)

Paradis et al. 2009 use a combination of corpus-driven and experimental methods to investigate their hypothesis about antonym canonicity. The starting point for their research is similar to that for part of the present research: the theoretical division of scholars between what they call the ‘lexical categorical model’ (e.g. Gross & Miller 1990) and the ‘cognitive prototype model’ (cf. among many others Herrmann et al. 1986, Murphy & Andrew 1993, Charles et al. 1994, Murphy 2003, Jones 2007, Paradis et al. 2009) i.e. the question whether antonymy is a lexical relation, which would mean that representations of opposite pairs are stored in the mental lexicon, or a conceptual relation, in which case opposite pairs would be computed ‘online’ in any given situation.

Paradis et al. (2009: 382) furthermore tackle the ‘chicken and egg’ question highlighted by Murphy & Andrew 1993 of whether lexical entrenchment is the consequence of the conceptual relation or whether the conceptual relation is strengthened by frequent co-occurrence of the two lexical items which form an antonym pair. The main hypothesis put forward is that ‘there is a limited core of highly opposable couplings that are strongly entrenched as pairs in memory and conventionalised as pairs in text and discourse, while all other couplings form a scale from more to less strongly related’ (Paradis et al. 2009: 381). This hypothesis seems to be confirmed by the data collected despite the fact that the different methods used do not always result in entirely compatible. However, it seems clear that the elicitation task used by Paradis et al. (2009: 414) shows that antonym canonicity is indeed a gradient phenomenon which manifests itself in the category-internal prototypicality structure of lexical opposition. Their judgement task shows 'that there is in fact a significant difference between canonical antonyms and other more contextual and less conventionalised pairings' (2009: 414).

As will be clear from the previous chapters, this is a similar approach to the present research which has taken great care to emphasise that the entrenchment of the lexical aspect of the antonymic relation alone cannot be enough to explain the range and strength of the phenomena observed in the data (cf. also Paradis et al. 2009: 386 and Murphy & Andrew 1993: 318). The study by Paradis et al. is the first one of this kind and comparing their results to the data collected in this study will allow for interesting conclusions to be drawn and may further substantiate both their hypothesis and the validity of the current approach.

Despite the overall similarities in the theoretical and practical approach, there are some significant differences between the present research and that of Paradis et al. which mainly stem from a different premise at the outset. The present research was designed to investigate the entirety of the category of lexical opposition in general whereas Paradis et al. focus on the border between canonical and non-canonical opposition of gradable adjectival antonym pairs in particular. Furthermore, there are differences between methods used by Paradis et al. 2009 and those the present approach (cf. Table 6.2 which lists the methods and measures used by both studies) as well as the fact that the comparative aspect (Chapter 5) has added an additional dimension to the present research which allows for direct comparison of the same concepts in two different languages.

**Table 6.2** *Comparison of empirical methods with Paradis et al. 2009*

| | CORPUS | JUDGEMENT | EXPERIMENTAL |
|---|---|---|---|
| present study | co-occurrence data from BNC<br>**M:*** t-score | untimed judgement task (web-based)<br>**M:** 7-point scale | behavioural LD task<br>**M:** RT |
| Paradis et al. 2009<br><br>*M = measure | co-occurrence data from BNC<br>**M:** expected co-occurrence | 'online' judgement task<br>**M:** 11-point scale<br>RT (but self-paced) | elicitation task<br>**M:** number of elicited lexemes |

Even though the empirical methods are different, there are enough similarities to make a comparison worthwhile. The following section will introduce the data from the present research in relation to the hypothesis presented by Paradis et al. 2009 and will then compare some canonical antonyms directly with the data from their study.

### 6.3.2 Antonymy – a prototype category?

The present study suggests that antonymy behaves like any other prototype category, with its members distributed along a continuum depending on the strength of membership in the category (cf. Lehrer 1990, Grandy 1992, Ungerer & Schmid 2006). According to Paradis et al. (2009: 381), however, there exist a small number of opposite pairs which are ‘special’ due to the additional strengthening of the opposition through a significant component of lexical association and entrenchment. This section will consider the data collected in the present study to see whether the findings presented by Paradis et al. can also be seen in this different dataset.

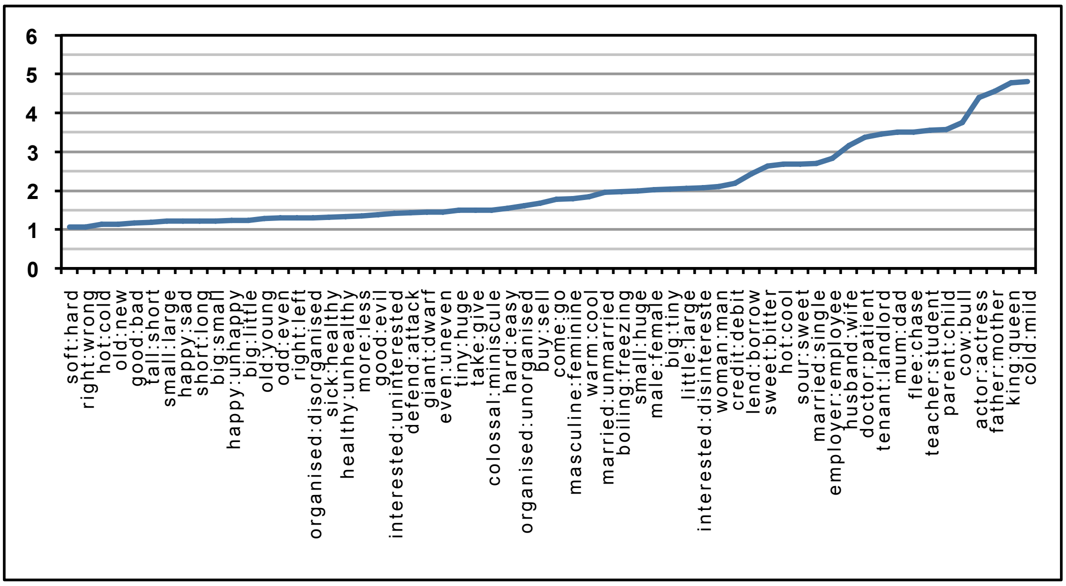



**Figure 6.2** *Antonym pairs by GOE-rating (Experiment 1)*

Figure 6.2 shows the results of the judgement tasks for all pairs which were included in Experiment 1 (for a graph of all pairs in the GOE-rating task, see Appendix 12). In the experiment there was a slight bias towards the lower end of the scale, which explains the initial ‘flat’ progression of the graph. The canonical opposite pairs which were part of the stimuli in the experiment are all in the leftmost quarter of the graph up to the last truly canonical pair (*sick:healthy*). After that, the graph climbs more or less steadily. There is no point in the graph at which a cut-off point could be declared since the increase is steady and consistent and does not display any clear breaks. The flat beginning of the graph could, if the data were displayed in a different way, prove to be very similar to data collected by Paradis et al. 2009 in their judgement task. If the antonym pairs were grouped into canonical and non-canonical antonyms, synonyms and other related pairs, there would be a significant difference between the first two categories as well as between the first and second and the third and fourth. This would result in a bar chart which is very similar to that in Paradis et al. (2009: 399) which led them to the conclusion that there is a set of canonical opposites which is different from the rest of the category and scores significantly higher in judgement tasks. However, since the purpose of the present study was not to investigate the boundary between canonical and non-canonical opposition but to look at the category of antonymy as a whole, the label 'canonical' antonym would now be given post-hoc and would thus be guided by the experimental results. Nevertheless, by Paradis et al.'s definition there are some clear-cut examples of canonical antonyms, in the data, namely those that form the base pairs for the clusters analysed in Chapter 3 as well as some additional canonical pairs which were included (cf. Figure 3.2 – for a full list of all pairs used for Figure 6.3 see Appendix 13). Therefore, Figure 6.3 below shows the two bar charts in comparison. Note that in Chart 1, ***eleven*** is the score given for the **most** antonymous pairs, whereas in Chart 2 the **most** antonymous pairs were given a score of ***one***.

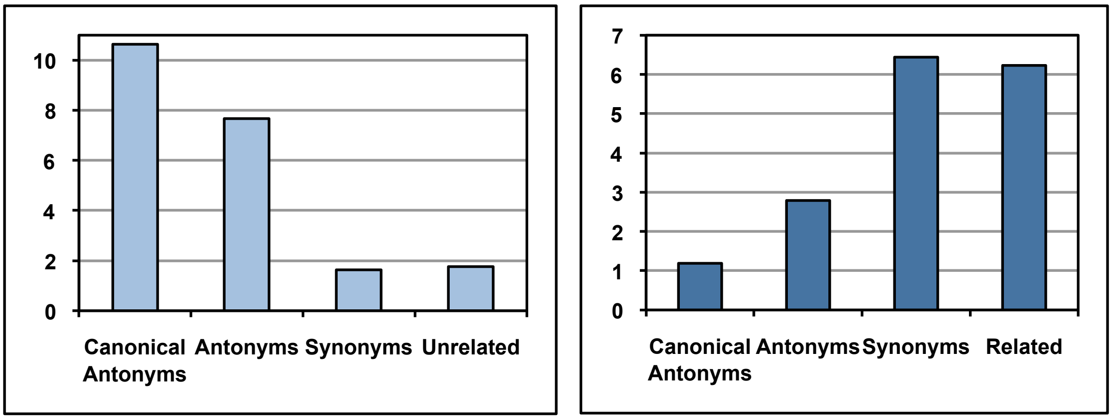



Chart 1: Paradis et al. 2009: 399

Chart 2: present study

**Figure 6.3** *Comparison of canonical and non-canonical pairs from Paradis et al. 2009 and the present study*

It can be seen from the charts in Figure 6.3 that, despite the different scales and the difference in methodology, the results of the GOE-rating method used in this study[2] are almost identical to those obtained by Paradis et al. 2009 in their computerised judgement task. The significant differences between the categories of *canonical antonyms* and *antonyms* as well as between those two and the remaining two categories (*synonyms* and *unrelated/related* lexemes) remain constant, as does the fact that the synonyms score slightly lower in both studies than the related/unrelated pairs. An interesting observation is that Paradis et al. used completely unrelated

[2] Only a selection of the pairs from the GOE-rating was used to match those in Paradis et al.'s study in terms of antonym/non-antonym type.

pairs for their fourth category (e.g. *big:white* and *heroic:young*) whereas the present work used associatively related lexemes which were neither synonymous nor antonymous (e.g. *tired:sick*, *scream:kick, book:page*). In addition to this difference, Paradis et al. confined their stimuli to adjectival pairs whereas my study used pairs from all major word classes in the fourth category. Nevertheless, the results are largely similar, which illustrates how the different presentation of the data and the assumption of certain categories guide our interpretation of the data (cf. also Paradis et al. 2009: 414). However, in this case, the graph in Figure 6.4 also shows the proposed group of canonical antonyms to be of a particularly homogeneous nature and seems to provide solid support to the identification of antonymy as a prototype category with a highly prototypical centre point made up of canonical antonyms.

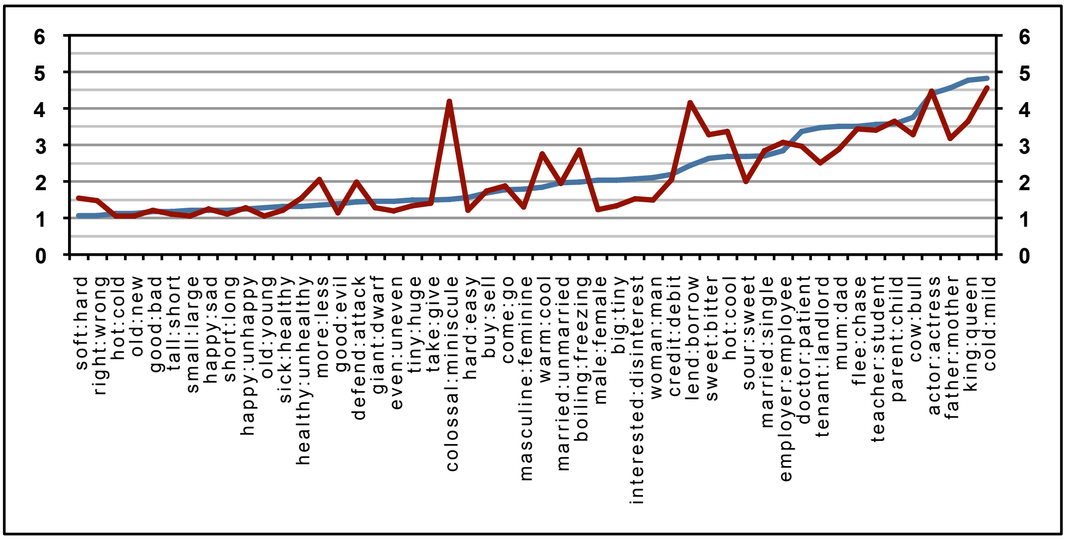



**Figure 6.4** *Comparison of GOE-rating English (blue) and German (red)*

The graphs in Figure 6.4 compare the German and English GOE-rating data. It would be easy to be led to the assumption that there are great differences between the English and German data by the rather erratic nature of the red graph (German GOE-rating data), but this is not at all the case. Figure 6.5 includes all pairs which were used as stimuli in Experiment 1 apart from those which do not have a direct German equivalent (e.g. *organised:disorganised*). It also includes those cases which were introduced in Chapter 5 (cf. 5.1.3) as displaying significant differences in the GOE-rating in the two languages. These cases, for example *lend:borrow* and *ausleihen:verleihen*, are on the whole easily explained and account for the jagged nature of the red graph above. The overall tendency of the graph is, however, remarkably similar to that of the English data and once again it does not show a clear point at which there is a sudden decrease in antonymic strength (which would result in a sudden rise in the graph).

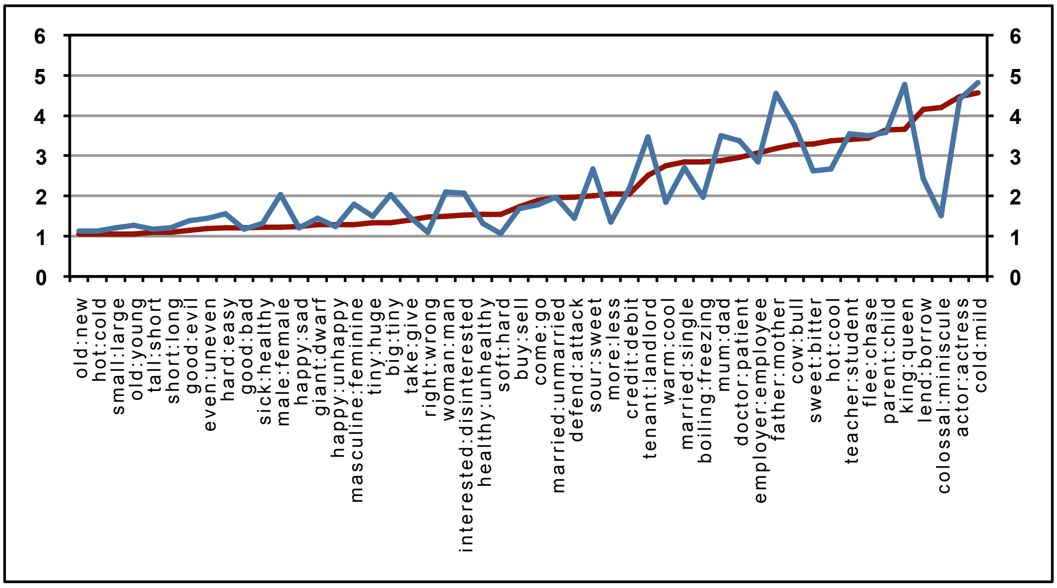



**Figure 6.5** *Data sorted by German GOE-rating*

Figure 6.5 presents the graph sorted by German GOE-results (red), and the red graph now displays a very similar progression to the blue graph in Figure 6.4 including the 'flat' beginning and then the consistent decrease of antonymic strength observed in the previous two figures. It seems that, as already stated in Chapter 5 (cf. 5.5), the distribution of antonyms along the canonicity continuum is cross-linguistically extremely similar.

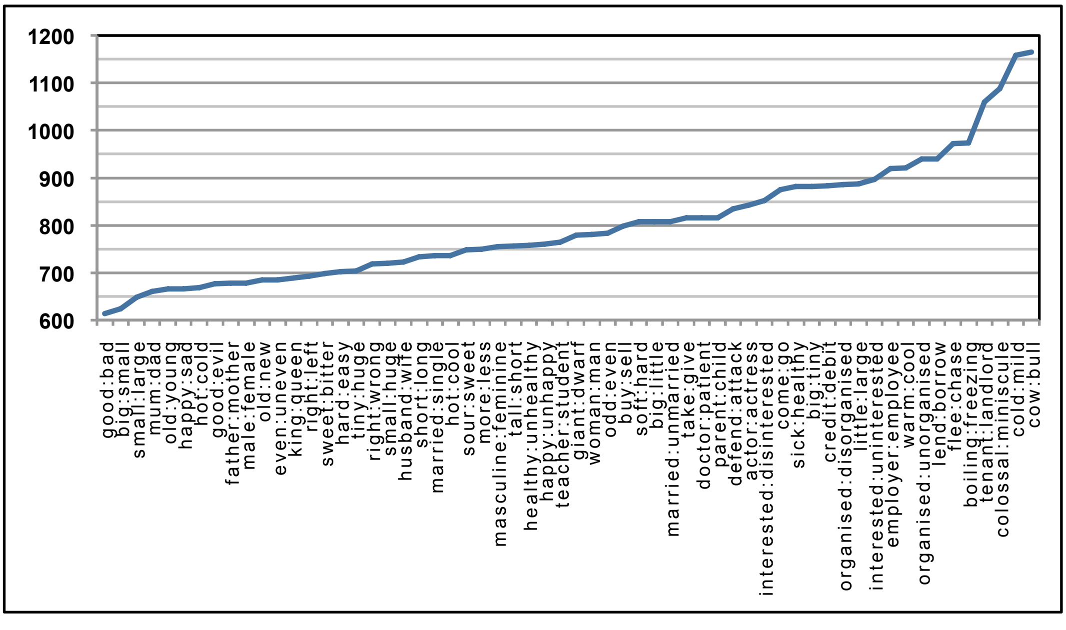



**Figure 6.6** *Antonym pairs (Experiment 1) by meanRT (in ms)*

Paradis et al.'s elicitation experiment and the behavioural lexical decision tasks used in the present work are more difficult to compare since they result in very different output data. However, the conclusions drawn by Paradis et al. that the elicitation experiments show the 'internal prototypicality structure of the category antonymy' (2009: 414) can be compared to the conclusions drawn in the present study on the basis of meanRT and standard deviation. An overview of this data as well as a more detailed discussion of certain aspects has already been given in Chapter 4 and will not be reiterated. The graph above (Figure 6.6) shows only the mean reaction time of all opposite pairs included in Experiment 1.

What is most striking in the graph above is the fact that, apart from the last four antonym pairs, the increase in reaction time is surprisingly linear. There are no noteworthy breaks or sudden increases. This very strongly points towards similar conclusions to those drawn by Paradis et al. 2009 and supports a construal of antonymy as a regular prototype category with its members evenly distributed along the canonicity continuum. Furthermore, the graph shows the expected effect of item length with the longer items clustered in the second half of the graph but this is by no means an effect which detracts from the overall picture. Another point to note is the comparatively steep increase in the first four or five pairs, which can be considered very strongly canonical (with the exception of *mum:dad*). The fast reaction times to some of the nominal pairs in the gender cluster (e.g. *mother:father* or *king:queen*) are unexpected since the GOE-rating results of those items were surprising (cf. 3.4.3 and 4.3.5.1) and lead to the assumption that co-occurrence plays a greater role in decisions made under pressure since those items may well be stored as a whole but not for reasons of antonymy. This observation could have consequences for the manner in which we think about lexical storage and activation.

On a comparative note, the German results also support the theory of a continuum of canonicity in both the GOE-rating and the behavioural task. In the figure below (Figure 6.7), a distribution can be seen which is very similar to that found in the English experiment. Simply looking at the graph overall, and not at the question of which opposite pair occupies which part of the reaction time continuum above, there is a visible increase (of about 30ms) after the first fifth of the scale. Some of the pairs in this first fifth are those which would be called very strongly canonical. However, to be able to draw firm conclusions, both GOE-rating and meanRT will need to be considered together.

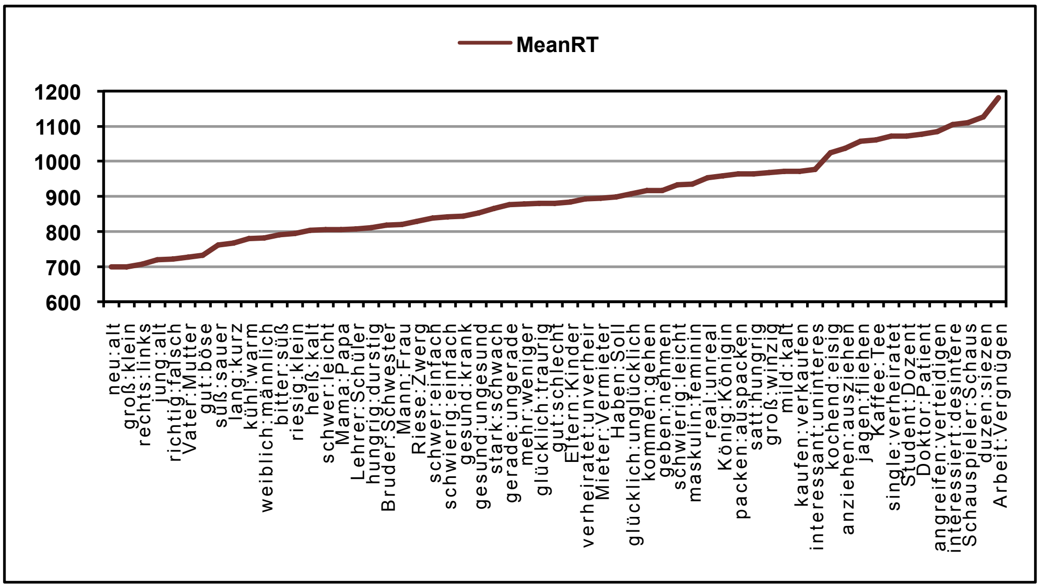



**Figure 6.7** *German antonym pairs by meanRT (in ms)*

The last figure (6.8) in this section shows both measures (GOE-rating and MeanRT) for the English canonical antonyms used in Experiment 1 ordered by MeanRT only. It is striking to see how different the two graphs are: the dark line, representing the behavioural data, is very similar to the other graphs discussed in this section and displays a steady increase in reaction time albeit over a smaller overall difference.

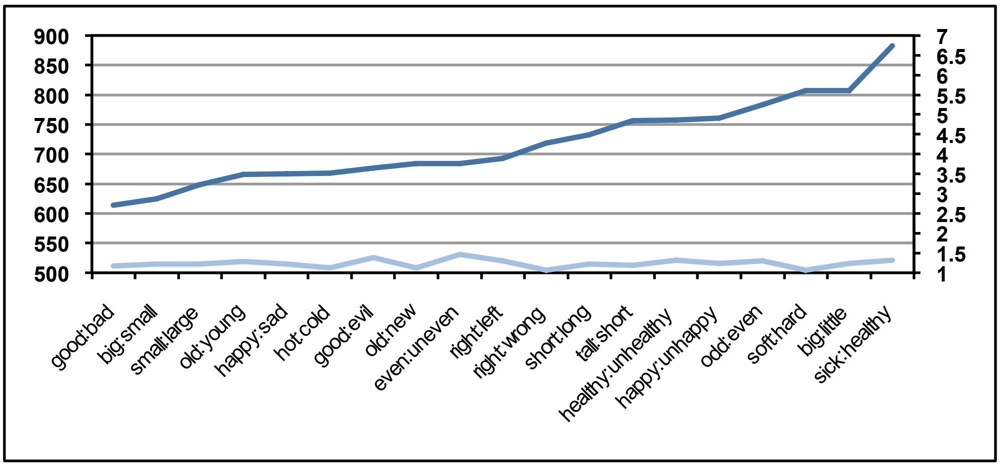



**Figure 6.8** *Canonical antonyms by MeanRT (in ms – dark) and GOE-rating (light)*

The light graph which represents the GOE-rating data on the other hand shows no such increase. There is hardly any difference between the pairs in terms of GOE-rating results. This finding lends support to Paradis et al.'s hypothesis of the peculiar make-up of the core of the category of antonymy. However, this is only supported by the judgement data and not by the results of the behavioural tasks. It remains to be seen whether there is an alternative explanation for either of the two measures, the most likely of which is the hypothesis that associative strength plays a greater role in the heavily timed behavioural task and in the GOE-rating task. This is a reasonable explanation, especially comparing the length of overall reaction times of the LD task in the present study and those in the self-paced judgement task conducted by Paradis et al. 2009 (between 600 and 1200ms in the former and 4000 and 6500ms in the latter).

To conclude the comparison between the two studies, it is clear that antonymy is a cognitive phenomenon and should be represented as such. Most antonym pairs are computed for any given context and these pairs pattern along a continuum of conventionality and antonymic strength.

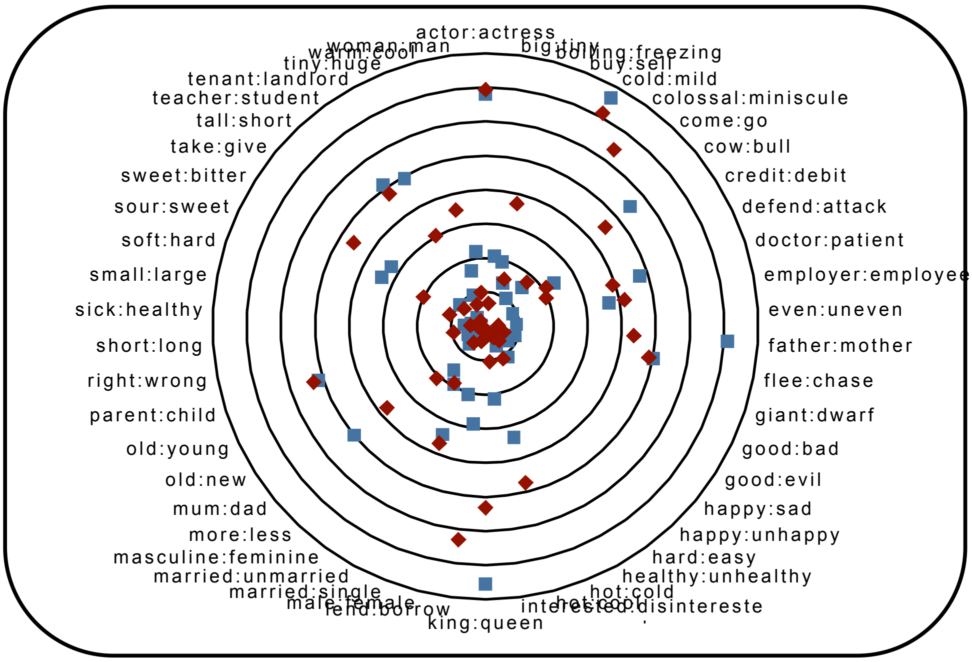



**Figure 6.9** *Antonymy as a prototype category*

There are a select few members of the category of antonymy for which this online computation does not apply – the canonical antonyms. However, there are probably fewer 'real' canonical opposites than assumed by, for example, Gross et al. 1990 who bisected the category of antonyms into canonical and non-canonical opposites. Theirs is not a view which can be upheld in the light of recent investigations into antonymy since even some of the canonical opposites which display somewhat different experimental results from the remaining category members show the same tendency to pattern along a continuum. If the category of antonymy were to be presented like early prototype categories by Aitchison (1994) and others, Figure 6.9 should be an accurate representation of the prototype category of English (blue) and German (red) antonyms.[3]

[3] Only those antonym pairs which were included in Experiment 1 are displayed in this figure.

# 7. A re-evaluation of antonym research

The *opposite* of *opposite*?
That's much too difficult, I quit.
(Wilbur 2004: 513)

This chapter will re-evaluate the notion of oppositeness in the light of the evidence collected and the hypothesis and arguments presented in the previous chapters, using the theoretical starting point(s) outlined in Chapter 2 as a foundation. The value of a multi-method approach and of the individual methodologies used will also be considered, before a summary of the results of the present study as well as an outline of suggested future avenues of investigation in the field of antonymy is presented.

## 7.1 What is *antonymy*?

From the very first mention and definition of the term *antonym,* provided by C.J. Smith in 1867, the term has undergone many re-definitions. It originated as an alternative to *counterterm* and was defined simply as 'a term which is the opposite or antithesis of another, a counter-term' (OED 1989). In Lyons' (1977) and Cruse's (1986) definitions, *antonym* came to refer to only a sub-group of opposites defined by certain specific properties (adjectival, gradable opposite pairs). When the notion of antonymy began to be investigated from a cognitive perspective, the term was used to refer to the relationship between two antonyms and its lexical, semantic and conceptual properties.

The question, raised in Cruse 1986, of which phenomena should be included in the category of opposition is one which has been considered repeatedly and has also formed a part of the present research. The question here is not only whether the category of converses, which shows different patterns from other types of opposition in the empirical studies, should be considered part of the category of lexical opposition but also whether any pairs beyond those central 'canonical' opposites are included in this category.

The definition provided by Jones in his corpus-driven investigation of antonymy seems to indicate that only those pairs qualify for antonym status which have become 'lexically "enshrined"' as opposites through frequent co-occurrence in particular environments. However, as the evidence above has shown, not all opposite pairs have this additional factor of associative strength as described in his definition. Thus, the question arising for any definition of antonymy is which opposites should be included in this category since this will inevitably determine the definition.

From the outset, the present research has made a case for antonymy as a category which includes more than simply gradable adjectival opposite pairs, and the notion of antonymy as a prototype category (cf. 6.3) was introduced early on in the study. Empirical evidence has confirmed the gradual decrease in antonymic strength from central, canonical opposites at one end of the continuum to pairs which depend strongly on contextual support in order to be judged as antonymic (e.g. *tea:coffee*, *credit:debit*) as their conceptual opposition is less salient. It therefore seems appropriate to define antonymy or lexical opposition as a cognitive relation of opposed concepts which are encoded by lexical items which, in the case of the most strongly opposed concepts, develop a very strong associative relation facilitating the entrenchment of the two lexemes as a pair and adding a lexical component to an essentially conceptual relation.

This definition includes all opposites discussed in this study, from those pairs which were ranked highest in the GOE-rating (e.g. *false:true*, *up:down*) and obtained the fastest reaction time in the behavioural experiments (*good:bad*) to those which were considered barely antonymic in the judgement task (e.g. *hungry:thirsty*, *dry:sweet*) to the pairs with the slowest reaction times (e.g. *cow:bull*). It further allows for the inclusion of antonyms from all lexical categories, since a native speaker's implicit definition of opposites is certainly not restricted to adjectival antonyms and also considers converses part of the category of opposites despite the difference in the type of the antonymic relation. This definition, and with it the approach outlined in Chapter 6, further goes in the direction of a request by Murphy (2003: 11f.) which is introduced in 2.3.2 and pleads for a theory of semantic relations which takes native speakers' judgements about the degree of the relation into account. It provides an inclusive approach to antonymy based on the conceptual opposition which nevertheless recognises and incorporates the importance of associative strength in antonym judgements.

## 7.2 A multi-method approach – advantages and disadvantages

The impetus for the multi-method approach taken in the present research stemmed from the observation that there were a large number of studies of antonymy which used different methodologies but a lack of investigations making use of more than one method. This section will consider the advantages and potential drawbacks of the present approach and highlight the insights gained by merging the different strands of data.

The first point which is important to take into account is the fact that the different strands of data provided three very different measures: t-scores from the corpus-based co-occurrence analysis of the BNC and the corpus of the IDS Mannheim, GOE-ratings from the judgement task and reaction times and error rates from the behavioural experiment. The comparison of all measures as well as some of the measures between English and German had to be considered very carefully to ensure the measures were in fact comparable. An effort was made when comparing German and English t-scores to ensure only the relative ranking of pairs by their t-scores was compared since the German corpus was significantly bigger than the BNC and thus the t-scores were correspondingly higher for the German data. However, as long as it is ensured that the measures being compared are related to each other in such a way as not to exaggerate certain effects, the benefits of this approach far outweigh this possible drawback.

The question of what can be gained from the combination of several methods should be relatively easily answered: a deeper and more thorough understanding of the phenomenon under investigation. The relation of antonymy had been considered from a number of different perspectives (cf. 2) ranging from highly theoretical and introspective classifications (Lyons 1977, Cruse 1986) to completely data-driven usage-based corpus studies (Jones 2002, Paradis et al. 2007). Nevertheless, there were never more than two methods used in any one study and mostly, the second method was only used for the selection of stimuli (cf. Murphy & Andrew 1993, Charles et al. 1994) for the purpose of ranking them to determine their semantic distance which allowed for a more accurate grouping for the 'main' empirical investigation. The present study, parallel to that of Paradis et al. 2009 discussed in 6.3, uses a new multi-method approach and it is the first of its kind to take into account comparative of such a large selection of opposites.

Bringing the three strands of data together has provided useful anchor points for the investigation of antonym canonicity and the factors which influence speaker judgements. The strategy of using discrepancies in the data as starting points for closer investigation and building data-driven hypotheses which could then be tested on the next set of data has proven effective, has resulted in a number of instances where newly introduced hypotheses could be substantiated by the new body of data and has allowed for the corroboration of theoretical approaches proposed by others working in the field (e.g. Paradis et al. 2009).

One further theoretically interesting point which has arisen in the course of the comparison of the results of the different methodological approaches, is the question of the extent to which the chosen methodology guides the researcher's interpretation of the data. It has become clear that some of the factors involved in the assessment of antonyms have different degrees of influence in the various types of methods used. A prime example of this is the fact that in the behavioural data collected for this study, associative strength had a much stronger influence than in the judgement task, especially at the top of the category. This held true for both languages and, if the interpretation of the data had only been based on the behavioural task, might have led to quite different results since it would then have been impossible to tell whether co-occurrence influences antonymic strength or associative strength (or both but to differing degrees). These confounding effects are more easily detected in a multi-method approach since a higher correlation between the measures is required to claim a relation between two factors or to assess the influence of a single factor.

## 7.3 Questions – asked and answered

The questions investigated in the present research were of two kinds: methodological and theoretical. On the methodological side, the multi-method approach and its merits have already been evaluated in the previous section and will thus not be discussed here at length. However, one of the greatest advantages of a multi-method approach is the fact that it is much more likely that idiosyncrasies in the data will be detected as such and investigated more thoroughly when they might otherwise have been misinterpreted as influential factors in their own right since there is no indication in the data from a single empirical method which would lead to a closer investigation of the relationship between particular factors. For example, without the information provided by the GOE-rating results, it would have been impossible to discern whether the fast reaction times for the nominal pairs like *mother:father* were due to antonymic or associative strength. Furthermore, associative strength and antonymic strength would have been assumed to correlate much more closely than is the case when the results of the judgement task are taken into account simply because the associative relation between the two members of a pair plays a more important role in the type of task used in the behavioural study. This, of course, leads to the question whether this method is then an appropriate tool for measuring semantic and conceptual rather than associative relations (cf. also Harley 2008: 179). However, as long as care is taken to ensure confounding factors (such as the associative relation in this case) are accounted for, these methods still provide useful evidence which advances our understanding of semantic relations.

On the theoretical side, three strongly connected questions were asked at the outset of the present research and each of them has been addressed throughout the evaluation of the evidence as well as in the theoretical discussion in 6. 3. These three overarching questions will all be addressed in turn here to discern to what extent the data introduced in this study has been able to provide the evidence necessary to answer them. Aside from the main aims of the investigation, various smaller questions which have been generated by the data-driven approach have been posed and investigated throughout the research. Some of the results pertaining to those questions will be summarised in the last part of this section.

(1) Is antonymy a gradient phenomenon or is there a clear distinction between canonical and non-canonical opposites?

There has been much discussion in the antonymy literature as to whether opposites can be split into two distinct categories of canonical and peripheral or direct and indirect opposites (cf. 2.4.1, 2.4.2 & 2.6) or whether there is a continuum of antonym canonicity along which all antonym pairs are distributed (cf. 2.6 & 6.3) depending on their antonymic strength and, potentially, other factors. The experimental evidence gathered for this study lends very strong support to a continuum approach according to which the category of antonymy is a prototype category with more and less central members and a gradual decrease in canonicity rather than a clear cut-off point between those two 'groups'. Paradis et al. (2009) found that there was a group of opposite pairs at the very top of the scale which seemed to be distinct from the rest of the group, which followed a very linear distribution. They claim that the pairs in this group, namely the classical canonical antonym pairs, are in addition to their high antonymic strength further facilitated by their associative strength. The data from the present study has shown a tendency towards the same phenomenon which, however, may not have been as clearly visible here since the stimuli were selected by very different criteria or the decrease in associative strength is, once again, gradual and thus the cut-off point once again can not be determined with any certainty. Further investigation is necessary to determine precisely how much associative strength influences antonym judgements.

(2) Which factors determine where on the antonym continuum a lexeme will be placed?

Following on from the discussion above, it is clear that antonym canonicity overall is determined not only by antonymic strength (which in itself is determined by a number of factors) but also by associative strength to some degree. A clear difference has to be made between antonym canonicity and antonymic strength: the former is how canonical an antonym pair is considered and it includes factors such as entrenchment through co-occurrence and morphological relatedness, while antonymic strength is only the conceptual aspect of the relation and includes factors such as inherent binarity, symmetry, overlap in semantic range and salience of antonymic features. There are a number of factors which play an important role in determining the place of a pair of opposites on the continuum and each of these has been discussed in detail in the present work. The interaction between these criteria in the cognitive model above is what finally determines the 'goodness' of an opposite pair. This is then modulated by the context as well as influenced by the associative lexical factors mentioned above. An investigation of whether computational modelling may be able to aid in the prediction of antonym canonicity by using the raw data collected in this study is one possible further means of exploiting the data.

(3) Is antonymy a relation between concepts or between the lexemes which encode these concepts?

One of the most important questions this study aimed to address was that of the nature of the relation of antonymy, which has previously given rise to much debate in the field (cf. 2.4.2.2). The analyses in this study have all provided evidence for a strong conceptual element in antonymy, which is primary, and in addition to that a secondary associative (lexical) element. Antonymy is, in the first place, a conceptual relation between concepts which are encoded by certain lexemes. This has been shown most strongly by the contrastive data where the same concepts obtain very similar ratings in the judgement tasks as well as reaction times which are very close to those for the equivalent pairs in the other language. The conceptual relation and the large shared semantic content of these concepts as well as their relationship with the underlying cognitive models, which, in order for the pair to be a good antonym pair, will be very similar, then provide the opportunity for the two lexemes to occur frequently in the same contexts.

This leads to the distribution seen in the prototype category above (cf. Figure 6.9) and the phenomenon observed by many other researchers of the very strong relation between canonical opposites, since the strength of their relationship is determined by two distinct types of criteria, one lexical and one conceptual.

The question of how opposites are stored in the mental lexicon is not as easily dealt with. Aside from the fact that where lexical storage is concerned, even the best supported arguments can only be hypotheses, it seems that lexical opposites are stored like other close associates and the entrenchment of the connection between the two lexemes and thus the degree to which they activate each other is determined by both antonymic and associative strength. The data collected in Experiment 2 and the analysis of the different relations between the four members of the English and German quartet pairs has revealed that some lexemes show a definite preference for one of their possible partners (e.g. *big:small* is better than *big:little*) which is not determined by the difference in associative strength. The difference was particularly apparent in the priming design used in Experiment 2 and was explained as inhibitory priming, which would lead to the assumption that the connection between *big:small* is stronger and blocks *little* as a possible second member of the antonym pair despite the fact that the two pairs scored almost identically in the judgement task. Many of the pairs, however, did not show any significant difference between the two possible combinations and produced almost identical reaction times (cf. 4.4.2), which suggests that both possible partners are activated when the first member of an antonym pair is given, provided they are of roughly equivalent antonymic and associative strength. The question of the asymmetry indicated by some of the results will have to be investigated further.

## 7.4 Future antonym research

This project has revealed several areas which would benefit from further research to extend and deepen our understanding of antonymy as a conceptual relation. An approach which included several different types of opposites as well as pairs from different lexical categories provided the broad canvas necessary for an investigation such as the one carried out here. However, to address the finer points of the phenomena observed in the study, it has been pointed out repeatedly that separate investigations should be carried out with sets of stimuli which are carefully chosen to address one very specific question rather than to provide the broad overview which was necessary to answer the questions asked in this study.

It would furthermore be of great interest to conduct an ERP study either with the same or with similar stimuli to investigate the effects of associative and semantic processing further. This may add to our understanding of the feature-matching process which seems to be involved in antonym processing and could shed light on the precise impact associative strength has on antonym judgements in the context of a behavioural lexical decision task.[1]

The cross-linguistic data has proved to be of great interest and it would be interesting to recreate the judgement task for languages which differ more from English than German does, to see whether the difference in linguistic structure and the cultural aspect has an effect on antonym judgements. This would allow for the comparison of the conceptual structure of certain concepts in several languages, especially if the judgement task is accompanied by an attribute listing task to establish the salient criteria for each of the concepts involved.

Two further areas which may be of interest to investigate but which are less closely related to the focus of the current study are the developmental and diachronic aspect of opposition which could include an investigation of antagonyms.[2] The developmental angle of antonymy has been investigated by several scholars (e.g. Clark 1972, Jones & Murphy 2005) and has provided support for the centrality of opposition within a child's lexicon and conceptual structure as well as for the idea that opposition constitutes a dynamic cognitive structuring principle. In light of the paradigmatic/syntagmatic shift (cf. Nelson 1977), it may be of use to conduct empirical studies with children at and around the age where the shift takes place to determine whether this changes the way opposites are judged and whether before this shift the syntagmatic criterion of co-occurrence has as much influence as it seems to have in adult perception of antonymy.

The diachronic aspect would initially be of particular interest in the case of the quartets such as those investigated here (cf. 3.1.4.3 & 5.2.3.4) since their existence could very well be a sign of lexical transition. The case of the English size adjectives is a particularly interesting one since the origin of some of the members of the central pairs (*big*) is not entirely clear. This lexeme, nevertheless, has become one of the central size adjectives while other lexemes which covered a similar

[1] A study of this kind is already being planned and will be conducted in the near future.
[2] Kotzor (submitted)

semantic range became obsolete (e.g. *mickle*). The investigation of the historical opposites would lend itself very well to a detailed discussion of a small number of pairs rather than a large-scale overview.

The present research has provided a large-scale overview of central and peripheral phenomena of opposition and used the differences between the canonical and less canonical opposites as starting points for a closer investigation of the lexical-semantic relation of antonymy.

# Bibliography

## Corpora and Databases

Kiss, G. R., C. Armstrong, R. Milroy & J. Piper (1973). An associative thesaurus of English and its computer analysis. In A. J. Aitken, R. W. Bailey and N. Hamilton-Smith (eds.) *The Computer and Literary Studies*. Edinburgh: University Press. [available at: http://www.eat.rl.ac.uk/ - last accessed March 2010].

*Korpora des Instituts für Deutsche Sprache*, (COSMAS – $\text{II}_{\text{web}}$) (2008). https://cosmas2.ids-mannheim.de/cosmas2-web/. Mannheim. [last accessed January 2010]

*The British National Corpus*, version 2 (BNC World) (2001). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. http://www.natcorp.ox.ac.uk. [last accessed April 2009]

Wilson, M.D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary (Version 2). *Behavioural Research Methods, Instruments and Computers* 20 (1), 6-11. [last accessed March 2009]

## Dictionaries

Hornby, A. S. (ed.) (2005). *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press [7th edition].

Kluge, F. (ed.) (2002). *Etymologisches Wörterbuch der deutschen Sprache*, (24th edition by E. Seebold), Berlin: Walter de Gruyter.

Langenscheidt (ed.) (2007). *Pocket Dictionary German*. Berlin/Munich: Langenscheidt KG.

Simpson, J. (ed.) (1989). *The Oxford English Dictionary*, Oxford: Oxford University Press [also OED online: www.oed.com - last accessed August 2010]

Summers, D. (ed.) (2003). *Longman Dictionary of Contemporary English*. Edinburgh: Longman – Pearson Education Ltd.

Summers, D. (ed.) (2002). Longman Language Activator. Harlow: Pearson Education.

## REFERENCES


Aitchison, J. (1994). *Words in the Mind. An introduction to the mental lexicon* (2nd edition). Oxford: Blackwell.

Aristotle (4th century BC). *De Interpretatione* and *Prior Analytics.* In J. Barnes (ed.) (1984). *The Complete Works of Aristotle*. Princeton: Princeton University Press.

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory and Cognition 8*, 439 - 512.

Carter, R. (1987). Is there a Core Vocabulary? Some implications for language teaching. *Applied Linguistics* 8, 178 - 193.

Cattell, J. M. (1947). On the time required for recognizing and naming letters and words, pictures and colors. In *James McKeen Cattell, Man of science* (vol. 1). Lancaster, PA: Science press, 13 - 25. [originally published 1888]

Charles, W. G. (2000). Contextual Correlates of Meaning. *Applied Psycholinguistics* 21, 505 - 524.

Charles, W. G., M. A. Reed & D. Derryberry (1994). Conceptual and associative processing of antonymy and synonymy. *Applied Psycholinguistics* 15, 329 - 354.

Charles, W. G. & G. A. Miller (1989). Contexts of antonymous adjectives. *Applied Psychology* 10, 357 - 375.

Clark, E. V. (1972). On the Child's Acquisition of Antonyms in two Semantic Fields. *Journal of Verbal Learning and Verbal Behaviour* 11, 750 - 758.

Collins, A. M. & E. F. Loftus (1975). A Spreading-Activation Theory of Semantic Processing. *Psychological Review* 82 (6), 407 - 428.

Croft, W. & D. A. Cruse (2004). *Cognitive Linguistics.* Cambridge: Cambridge University Press.

Cruse, D. A. (1976). Three classes of antonym in English. *Lingua* 38, 281 - 292.

Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Cruse, D. A. (1990). Prototype Theory and Lexical Semantics. In Tsohatzidis (1990), 382 - 402.

Cruse, D. A. (1992). Antonymy Revisited: Some thoughts on the relationship between words and concepts. In A. Lehrer & E. Kittay (eds.) *Frames, Fields and Contrasts.* , Hillsdale, NJ: Lawrence Erlbaum, 289 - 306.

Cruse, D. A. (1994). Prototype theory and lexical relations. *Rivista di Linguistica 6*, 167 - 188.

Cruse, D. A. (2000). *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.

Cruse, D. A. & P. Togia (1995). Towards a Cognitive Model of Antonymy. *Lexicology* 1, 113 - 141.

Cutler, A. (1981). The reliability of speech error data. *Linguistics* 19, 561 - 582.

Deese, J. (1964). The associative structure of some common English adjectives. *Journal of Verbal Learning and Verbal Behaviour* 3, 347 - 357.

Deese, J. (1965). *The structure of associations in language and thought.* Baltimore: The Johns Hopkins Press.

Eco, U. (2003). *Mouse or Rat? Translation as Negotiation*. London: Weidenfeld & Nicolson.

Ervin, S. M. (1961). Changes with age in the verbal determinants of word-association. *American Journal of Psychology* 74, 361 - 372.

Ervin, S. M. & W. Miller (1963). Language Development. *National Society for the Study of Education Yearbook*, 108 - 143.

Evans, V. & M. Green (2006). *Cognitive Linguistics: An introduction*. Edinburgh: Edinburgh University Press.

Fellbaum, C. (1995). Co-occurrence and Antonymy. *International Journal of Lexicography* 8, 281 - 303.

Fillmore, C. J. (1985). Frames and the Semantics of Understanding. *Quaderni di Semantica* 6, 222 - 254.

Foster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. C. T. Walker (eds.) *New approaches to language mechanisms*. Amsterdam: North Holland, 257 - 287.

Foster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (eds.) *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., 27 - 85.

Francis, W. N. & H. Kucera (1982). *Frequency analysis of English usage*. Boston, MA: Houghton Mifflin.

Fromkin, V. A. (1973). *Speech Errors as Linguistic Evidence.* The Hague: Mouton.

Gazzaniga, M. & G. Miller (1989). The Recognition of Antonymy by a Lanugage-Enriched Right Hemisphere. *Journal of Cognitive Neuroscience* 1, 187 - 193.

Grandy, R. E. (1992). Semantic Fields, Prototypes, and the Lexicon. In Lehrer & Kittay, 103 - 122.

Gross, D. & K. Miller. (1990). Adjectives in WordNet. *International Journal of Lexicography* 3, 265 - 277.

Gross, D., U. Fischer & G. A. Miller. (1988). Antonymy and the representation of adjectival meanings. *Cognitive Science Laboratory Report 13*. Princeton: Princeton University.

Gross, D., U. Fischer & G. A. Miller. (1989). The organization of adjectival meanings. *Journal of Memory and Language* 28, 92 - 106.

Harley, T. (2008). *The Psychology of Language. From Data to Theory*. New York: Psychology Press.

Herrmann, D. J., R, Chaffin, M. P. Daniel & R. S. Wool (1986). The role of elements of relation definition in antonym and synonym comprehension. *Zeitschrift für Psychologie* 194, 133 - 53.

Herrmann, D. J. & R. Chaffin (1986). Comprehension of semantic relations as a function of the definitions of relations. In F. Klix & H. Hagendorf (eds.) *Human Memory and Cognitive Capabilities*. Amsterdam: Elsevier, 311 - 319.

Hoey, M. (1991). *Patterns of Lexis in Text*. Oxford: Oxford University Press.

Horn, L. (1989). *A natural history of negation*. Chicago: The University of Chicago Press.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A *micro*analytic review. *Psychonomic Bulletin & Review* 10 (4), 785 - 813.

Jeon, H., K. Lee, Y. Kim & Z. Cho (2009). Neural substrates of semantic relationships: Common and distinct left-frontal activities for generation of synonyms vs. antonyms. *Neuroimage* 48, 449 - 457.

Jones, S. (2002). *Antonymy. A corpus-based apporach*. London: Routledge.

Jones S. & M.L. Murphy (2005). Using Corpora to Investigate Antonym Acquisition. *International Journal of Corpus Linguistics* 10 (3), 401 - 422.

Jones, S. (2006). A lexico-syntactic analysis of antonym co-occurrence in spoken English. *Text and Talk* 26 (2), 191 - 216.

Jones, S. (2007). Opposites in discourse: A comparison of antonym use across four domains. *Journal of Pragmatics* 39, 1105 - 1119.

Justeson, J. & S. Katz (1991). Co-occurrences of antonymous adjectives and their contexts. *Computational Linguistics* 17, 1 - 19.

Justeson, J. & S. Katz. (1992). Redefining Antonymy: The Textual Structure of a Semantic Relation. *Literary and Linguistic Computing* 7, 176 - 184.

Kjellmer, G. (2005). Negated Adjectives in Modern English. A corpus based study. *Studia Neophilologica* 77, 156 - 170.

Kneale, W. & M. Kneale (1962). *The Development of Logic*. Oxford: Oxford University Press.

Lakoff, G. (1987). *Women, fire and dangerous things. What categories reveal about the mind.* Chicago/London: University of Chicago Press.

Lakoff, G. & M. Johnson. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lehrer, A. (1985). Markedness and antonymy. *Linguistics* 21, 397 - 429.

Lehrer, A. (1990). Prototype theory and its implications for lexical analysis. In Tsohatzidis (1990), 368 - 381.

Lehrer, A. (2002). Gradable antonymy and complementarity. In D. A. Cruse, F. Hundsnurscher, M. Job & P. Lutzeier (eds.) *Handbook of Lexicology*, Berlin: de Gruyter, 498 - 506.

Lehrer, A. & K. Lehrer (1982). Antonymy. *Linguistics and Philosophy* 5, 483 - 501.

Lehrer, A. & E. Kittay (eds.) (1992). *Frames, fields and contrasts*. Hillsdale, NJ: Lawrence Erlbaum, 75 - 102.

Lutzeier, P. (1997). Gegensinn als besondere Form lexikalischer Ambiguität, *Linguistische Berichte* 171, 381 - 395.

Lutzeier, P. (2001). Polysemie mit spezieller Berücksichtigung des Gegensinns. *Lexicographica* 17, 69 - 91.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

Lyons, J. (1995). *Linguistic semantics: An introduction*. Cambridge: Cambridge University Press.

McClelland, J. L. & D. E. Rumelhardt (1981). An interactive activation model of context effects in letter perception: Part 1. An account of the basic findings. *Psychological Review* 88, 375 - 407.

McNamara, T. P. (1992). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18, 1173 - 1190.

McNamara, T. P. (1994). Theories of priming: II. Types of prime. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20, 507 - 520.

McNamara, T. P. & J. Altarriba (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language* 27, 545 - 559.

Mettinger, A. (1994). *Aspects of Semantic Opposition in English*. Oxford: Oxford University Press.

Miller, G. (1990). WordNet: An on-line lexical database. *International Journal of Lexicography* 3, 235 - 312.

Miller, K. J. (1998). Modifiers in WordNet. In C. Fellbaum (ed.). *WordNet: an electronic lexical database.* Cambridge, MA: The MIT Press, 47 - 68.

Miller, G. A. & C. Fellbaum (1991). Semantic networks of English. *Cognition* 41, 197 - 229.

Mohammad, S., B. Dorr & G. Hirst (2008). Computing Word-Pair Antonymy. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, 982 - 991.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review* 76, 165 - 178.

Morton, J. (1970). A functional model for human memory. In D. A. Norman (ed.) *Models of human memory*. New York: Academic Press, 203 - 260.

Morton, J. (1979). Facilitation in word recognition: Experiments causing change in the logogen model. In P. A. Kolers, M. E. Wrolstad & M. Bouma (eds.) *Processing of visible language*. New York: Plenum, 259 - 268.

Muehleisen, V. (1997). *Antonymy and Semantic Range in English*. Unpublished PhD Dissertation: Northwestern University.

Müller, W. (1998). *Das Gegenwort-Wörterbuch: ein Kontrastwörterbuch mit Gebrauchshinweisen*. Berlin/New York: de Gruyter.

Murphy, G. L. & J. M. Andrew (1993). The Conceptual Basis of Antonymy and Synonymy in Adjectives. *Journal of Memory and Language* 32, 301-319.

Murphy, L. M. (1993). *In opposition to an organised lexicon: pragmatic principles and lexcial semantic relations*, University of Illinois.

Murphy, M. L. (1994). A note on pragmatic markedness. *Pragmatics and Language Learning Monographs*, 277 - 288.

Murphy, M. L. (2000). Knowledge of words versus knowledge about words: the conceptual basis of lexical relations. In B. Peeters (ed.) *The lexicon - encyclopedia interface*. Amsterdam: Elsevier, 317 - 348

Murphy, M. L. (2003). *Semantic Relations and the Lexicon*. Cambridge: Cambridge University Press.

Murphy, M. L. (2006). Antonyms as lexical constructions: or, why paradigmatic construction isn't an oxymoron. In Schönefeld, D. (ed.) *Constructions all over: case studies and theoretical implications.* Special volume of Constructions, SV1-8/2006. http://www.constructions-online.de/. [accessed 19th January 2008]

Murphy, M. L., C. Paradis, C. Willners & S. Jones (2009). Discourse functions of antonymy: a cross-linguistic investigation. *Journal of Pragmatics* 41(11), 2159 - 2184.

Needham, R. (1987). *Counterpoints*. Berkeley: University of California Press.

Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: a review of research and theory. *Psychological Bulletin* 84, 93 - 116.

Ogden, C. K. (1967). *Opposition: a linguistic and psychological analysis*. Bloomington: Indiana University Press [originally published in 1932].

Osgood, C., G. J. Suci & P. H. Tannenbaum (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Pachella, R. G. (1974). The interpretation of reaction time in information processing research. In B. H. Kantowitz (ed.) *Human information processing: Tutorials in performance and cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., 41 - 82.

Palermo, D. S., & J. J. Jenkins (1964). *Word association norms: grade school through college*. Minneapolis: University of Minnesota Press.

Paradis, C., C. Willners, M.L. Murphy & S. Jones (2007). Googling for opposites: a web-based study of antonym canonicity. *Corpora* 2 (2), 129 - 155.

Paradis, C., C. Willners (2007). Antonyms in Dictionary Entries: Methodological Aspects. *Studia Linguistica* 61 (3), 261 - 277.

Paradis, C., C. Willners & S. Jones (2009). Good and bad opposites. Using textual and experimental techniques to measure antonym canonicity. *The Mental Lexicon* 4 (3), 380 - 429.

Paradis, C., C. Willners, S. Löhndorf & M.L. Murphy (2006). Quantifying aspects of antonym canonicity in English and Swedish: textual and experimental. Presented at *Quantitative Investigations in Theoretical Linguistics* 2, Osnabrück, Germany, 1–2 June.

Quillian, M. R. (1962). *A semantic coding technique for mechanical English paraphrasing*. Technical report, Mechanical translation Group, Research Laboratory of Electronics, M.I.T.

Quillian, M. R. (1967). Word concepts: A theory and simulation of some basic semantic capabilities. *Behavioural Science* 12 (5), 410 - 430.

Reetz, H., & A. Kleinmann (2003). Multi-subject hardware for experiment control and precise reaction time measurement. In M. J. Solé , D. Recasens, & J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences.* Barcelona, 1489 - 1492.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (ed.) *Cognitive development and the acquisition of language*. New York/London: Academic Press, 111 - 144.

Rosch, E. (1975). Cogntive representation of semantic categories. *Journal of Experimental Psychology* 104 (3), 193 - 233.

Rosch, E. (1978). Principles of Categorization. In E. Rosch & B. B. Lloyd (eds.) *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum, 27 - 48.

Rosch, E. & C. Mervis (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7 (4), 573 - 605.

Sabourin, L. & G. Libben (2000). Lexical Processing of Synonymy and Antonymy: An exploration of Task and Word Form Differences. In *Papers in Experimental and Theoretical Linguistics*. University of Alberta working papers 5, 114 - 136.

Sapir, E. (1944). Grading. *Philosophy of Science* 11, 83 - 116.

Saussure, F. (1916). *Cours de linguistique générale*. Published by C. Bally, A. Sechehaye & A. Riedlinger. (1965). Paris: Bayot.

Smith, C. J. (1967). *Synonyms and Antonyms*. London: Bell & Daldy.

Taylor, J. R. (1989). *Linguistic Categorization: Prototypes in linguistic theory*. Oxford: Clarendon Press.

Taylor, J. R. (1990). Schemas, prototypes, and models. In Tsohatzidis (1990), 521 - 534.

Tsohatzidis. S. L. (ed.) (1990). *Meanings and Prototypes*. London/New York: Routledge.

Ungerer, F. & H.-J. Schmid (2006). *An introduction to cognitive linguistics*. Edinburgh: Longman/Pearson Education.

Van Jaarsveld, H. J. & I. Draskovic (2003). Effects of collocational restrictions in the interpretation of adjective-noun combinations. *Language and Cognitive Processes* 18, 47 - 59.

Vandeloise, C. (1990). Representation, prototypes and centrality. In S.L Tsohatzidis (ed.) *Meaning and Prototypes*. London/New York: Routledge, 403 - 446.

Varley, R. (1991). Reference, sense and antonymy in the assessment of lexical semantic abilities in aphasia. *Aphasiology* 5, 149 - 170.

Wilbur, R. (2000). *Opposites, more opposites and a few differences*. New York: Harcourt.

Wilbur, R. (2004). *Collected Poems 1943-2004*. New York: Harcourt, Inc.

Zimmer, K. (1964). Affixal Negation in English and other languages: an investigation of restricted productivity. Supplement to *Word* 20 (2).

## **Appendix 1** Sample Goodness-of-exemplar-rating Questionnaire (English) Questionnaire I-1

Opposites are something we use very often in everyday language and to most of us there are some opposites which seem to be better than others.

For example, *maximize – minimize* is generally considered a very good opposite pair whereas *chair – table* is, if at all, only considered a fairly bad example of opposition.

The following task consists of a list of opposites which you are asked to judge on a scale from 1 to 7, 1 being a very good example of opposition, 7 being a very bad one and 4 indicating a moderately good opposite pair.

Please rely on your initial response and do not think too much about your answers. Once you have ranked an opposite pair, please do not go back to it afterwards to change the ranking.

Thank you very much for participating!

| | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|---|
| cold – friendly | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| fish – fowl | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| rain – wind | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| dry – sweet | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| tutee – tutor | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| fit – hot | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| work – play | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| glove – hand | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| helpful – helpless | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| misbehave – behave | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| soft – easy | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| excellent – bad | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| excellent - atrocious | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| blue – orange | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| occupier – owner | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| yell – call | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| short – squat | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| mad – angry | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| defend – attack | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| crown – king | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| good – disobedient | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| big – small | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| behind – in front | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| little – gigantic | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| tired – sick | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| new – old | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| mediocre – brilliant | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| hard – easy | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| dark – light | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| bad – satisfactory | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| incorrect – right | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| unhappy – happy | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| father – mother | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| sweltering – nippy | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| lukewarm – tepid | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| short – long | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| unorganised – organised | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| credit – debit | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| coffee – tea | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| dog – cat | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| out – in | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| spicy – mild | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| unhealthy – healthy | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| still – sparkling | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| present – future | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| animal – dog | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| suitable – wrong | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| tall – short | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| easy – difficult | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| little – big | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| loyal – disloyal | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| donate – steal | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| top – bottom | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| current – former | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| old – young | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| together – apart | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| trade – purchase | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| red – green | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| traditional – trendy | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| tiny – huge | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| even – uneven | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| ask – answer | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| interested – disinterested | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| single – married | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| answer – reply | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| father – daughter | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| colossal – miniscule | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| cold – distant | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| light – heavy | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| dirty – clean | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| disorganised – organised | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| tenant – landlord | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| awake – asleep | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| uncle – aunt | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| next to – opposite | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| away – home | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| soft – hard | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| walk – crawl | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| rent – let | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| lease – let | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| mentor – protégé | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| freezing – boiling | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| sober – drunk | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| brick – mortar | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| pack – unpack | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| pupil – teacher | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| escape – hunt | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| take – give | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| chilly – steaming | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| up – down | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| possible – impossible | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| last – first | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| uneasy – easy | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| egg – chicken | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| cold – hot | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| cold – mild | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| teacher – student | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| king – queen | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| similar – different | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| dead – fresh | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| sinful – virtuous | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| indirectly – directly | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| fast – slow | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| serious – trivial | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

If you have any comments or questions regarding the task please put them in the box below. Thank you.

### LINKS TO ONLINE QUESTIONNAIRES

ENGLISH

Questionnaire I-1 http://www.surveymonkey.com/s/6X5XYZX
Questionnaire I-2 http://www.surveymonkey.com/s/6NCJLPM
Questionnaire II-1 http://www.surveymonkey.com/s/6NWPS6D
Questionnaire II-2 http://www.surveymonkey.com/s/6NMDQFS

GERMAN

Questionnaire I-1 http://www.surveymonkey.com/s/3FS95D7
Questionnaire I-2 http://www.surveymonkey.com/s/3FR7BMR
Questionnaire II-1 http://www.surveymonkey.com/s/3FSR6NZ
Questionnaire II-2 http://www.surveymonkey.com/s/3FSHPV3

**Appendix 2** English and German word pairs included in the GOE-rating

| ENGLISH | | GERMAN | |
|---|---|---|---|
| **Word 1** | **Word 2** | **Word 1** | **Word 2** |
| absent | present | abwesend | anwesend |
| actress | actor | Schauspielerin | Schauspieler |
| aloof | amiable | verschlossen | unverschlossen |
| animal | dog | Tier | Hund |
| answer | reply | Antwort | Entgegnung |
| apprentice | master | Lehrling | Meister |
| ask | answer | fragen | antworten |
| awake | asleep | wach | schlafend |
| away | home | weg | daheim |
| bad | satisfactory | schlecht | zufriedenstellend |
| bad | good | schlecht | gut |
| badly | well | kaufen | verkaufen |
| behind | in front | hinter | vor |
| big | tiny | groß | winzig |
| big | small | groß | klein |
| bland | hot | fad | scharf |
| blue | orange | blau | orange |
| book | page | Buch | Seite |
| brick | mortar | unbunt | bunt |
| brief | long | Bein | Fuß |
| brother | sister | Bruder | Schwester |
| buy | spend | kaufen | ausgeben |
| buy | sell | kaufen | verkaufen |
| carrot | pea | Karotte | Erbse |
| child | parent | Kinder | Eltern |
| chilly | steaming | kühl | drückend |
| chilly | warm | frostig | warm |
| clean | tidy | sauber | ordentlich |
| clever | bright | clever | helle |
| coffee | tea | Kaffee | Tee |
| cold | friendly | kalt | freundlich |
| cold | distant | kalt | fremd |

| | | | |
|---|---|---|---|
| cold | mild | kalt | mild |
| cold | hot | kalt | heiß |
| cold | sticky | kalt | klebrig |
| colossal | miniscule | kolossal | unscheinbar |
| contemporary | ancient | offen | verschlossen |
| correct | mistaken | verlernen | lernen |
| cow | bull | Kuh | Bulle |
| credit | debit | Haben | Soll |
| crown | king | Krone | König |
| current | former | gegenwärtig | früher |
| dad | mum | Papa | Mama |
| damage | repair | beschädigen | reparieren |
| dark | light | dunkel | hell |
| daughter | mother | Tochter | Mutter |
| dead | fresh | tot | frisch |
| dead | gone | tot | weg |
| defend | attack | verteidigen | angreifen |
| dirty | clean | schmutzig | sauber |
| disorganised | organised | fest | weich |
| dog | cat | Hund | Katze |
| donate | steal | spenden | stehlen |
| double | single | doppelt | einfach |
| dry | sweet | trocken | süß |
| easy | uneasy | mild | lau |
| easy | difficult | einfach | schwierig |
| egg | chicken | Ei | Huhn |
| employer | employee | Chef | Angestellter |
| escape | hunt | ausbrechen | jagen |
| even | uneven | gerade | ungerade |
| excellent | atrocious | exzellent | miserabel |
| excellent | bad | exzellent | schlecht |
| exclude | include | ausschließen | miteinbeziehen |
| fair | poor | angemessen | mangelhaft |
| false | true | falsch | wahr |
| far | near | weit | nah |
| fast | slow | schnell | langsam |

| | | | |
|---|---|---|---|
| father | mother | Vater | Mutter |
| father | daughter | Vater | Tochter |
| figuratively | literally | einflusslos | einflussreich |
| finger | hand | Finger | Hand |
| finish | continue | Hahn | Henne |
| fish | fowl | Fisch | Fleisch |
| fit | hot | Katze | Maus |
| flammable | inflammable | einwilligen | befolgen |
| flee | chase | fliehen | jagen |
| freezing | boiling | eisig | tropisch |
| giant | dwarf | Riese | Zwerg |
| glove | hand | früh | spät |
| go | come | gehen | kommen |
| good | disobedient | brav | ungehorsam |
| good | mediocre | gut | mittelmäßig |
| good | evil | gut | böse |
| gracious | mean | satt | hungrig |
| happy | sad | glücklich | unglücklich |
| hard | easy | schwer | einfach |
| helpful | helpless | hilfreich | hilflos |
| helpless | helpful | hilflos | hilfreich |
| hot | cool | heiß | kühl |
| humid | hot | schmal | dick |
| hungry | thirsty | hungrig | durstig |
| imprecise | exact | unpräzise | exakt |
| inaccurate | accurate | inakkurat | akkurat |
| incorrect | right | inkorrekt | richtig |
| indirectly | directly | indirekt | direkt |
| interested | disinterested | interessiert | uninteressiert |
| interested | uninterested | trostlos | verzagt |
| keep | sell | behalten | verkaufen |
| king | queen | König | Königin |
| large | little | halb | ganz |
| last | first | letztes | erstes |
| lease | let | lehren | lernen |
| legal | illegal | legal | illegal |

| | | | |
|---|---|---|---|
| lend | borrow | verleihen | ausleihen |
| less | more | weniger | mehr |
| life | death | Leben | Tod |
| light | heavy | leicht | schwer |
| little | gigantic | klein | riesig |
| little | big | klein | groß |
| long | narrow | lang | schmal |
| loyal | disloyal | loyal | unloyal |
| lukewarm | tepid | Weib | Mann |
| mad | angry | verrückt | wütend |
| male | female | männlich | weiblich |
| married | unmarried | verheiratet | unverheiratet |
| married | divorced | verheiratet | geschieden |
| masculine | feminine | maskulin | feminin |
| mediocre | brilliant | mittelmäßig | brilliant |
| mentor | protégé | Tasse | Henkel |
| miniature | gigantic | mini | riesig |
| misbehave | behave | unmoralisch | moralisch |
| mobile | landline | Handy | Festnetz |
| mug | cup | Becher | Tasse |
| murderer | victim | Mörder | Opfer |
| nephew | niece | Neffe | Nichte |
| new | old | neu | alt |
| next to | opposite | neben | gegenüber |
| occupier | owner | Bewohner | Besitzer |
| odd | even | moralisch | amoralisch |
| old | young | alt | jung |
| out | in | hinaus | hinein |
| over | under | über | unter |
| pack | unpack | packen | auspacken |
| pale | dark | blass | dunkel |
| patient | doctor | Patient | Doktor |
| possible | impossible | möglich | unmöglich |
| present | future | Gegenwart | Zukunft |
| present | past | Gegenwart | Vergangenheit |
| pupil | teacher | Schüler | Lehrer |

| | | | |
|---|---|---|---|
| quiet | still | ruhig | still |
| rain | wind | Regen | Wind |
| rational | irrational | rational | irrational |
| red | green | rot | grün |
| rent | let | mieten | vermieten |
| right | left | rechts | links |
| right | wrong | richtig | falsch |
| rose | flower | Rose | Blume |
| sad | unhappy | traurig | unglücklich |
| scorching | glacial | amoralisch | unmoralisch |
| scream | kick | weich | fest |
| serious | trivial | scheußlich | verrucht |
| short | squat | bunt | einfarbig |
| short | polite | kurz | höflich |
| short | long | kurz | lang |
| sick | healthy | krank | gesund |
| similar | identical | ähnlich | identisch |
| similar | different | ähnlich | verschieden |
| sinful | virtuous | sündig | tugendhaft |
| single | married | single | verheiratet |
| small | huge | klein | gewaltig |
| small | tall | staatlich | privat |
| small | large | Fee | Hexe |
| small | large | schwach | stark |
| sober | drunk | nüchtern | betrunken |
| soft | easy | weich | einfach |
| soft | hard | weich | hart |
| sour | sweet | sauer | süß |
| spicy | mild | würzig | mild |
| square | long | viereckig | lang |
| still | sparkling | privat | dienstlich |
| succeed | fail | sommerlich | frostig |
| suitable | wrong | passend | falsch |
| sweet | bitter | süß | bitter |
| sweltering | nippy | formal | praktisch |
| take | give | nehmen | geben |

| | | | |
|---|---|---|---|
| tall | short | schwierig | leicht |
| teacher | student | Dozent | Student |
| tenant | landlord | Mieter | Vermieter |
| tiny | huge | winzig | riesig |
| tipsy | inebriated | irreal | real |
| tired | sick | Wasser | Feuer |
| together | apart | zusammen | alleine |
| tomorrow | today | morgen | heute |
| top | bottom | oben | unten |
| town | country | Stadt | Land |
| trade | purchase | handeln | erwerben |
| traditional | trendy | traditionell | modisch |
| tutee | tutor | siezen | duzen |
| uncle | aunt | Onkel | Tante |
| undress | dress | ausziehen | anziehen |
| uneasy | easy | alt | grau |
| unhappy | happy | unglücklich | glücklich |
| unhealthy | healthy | ungesund | gesund |
| unorganised | organised | erlauben | verbieten |
| unsuitable | suitable | unpassend | passend |
| untidy | tidy | unordentlich | ordentlich |
| up | down | hoch | runter |
| valley | hill | Tal | Berg |
| walk | crawl | gehen | krabbeln |
| warm | cool | warm | kühl |
| white | coloured | weiß | farbig |
| wife | husband | Frau | Mann |
| wintry | sultry | real | unreal |
| woman | man | wohin | woher |
| work | play | Arbeit | Vergnügen |
| wrong | bad | falsch | schlecht |
| yell | call | brüllen | rufen |
| yellow | blue | gelb | blau |
| | | | |
| | not translated | | |

**Appendix 3** T-scores and GOE-rating for all English pair

| Word 1 | Freq | Word 2 | Freq | Total FoC | T Score | GOE 1-2 | GOE 2-1 | GOE Overall |
|---|---|---|---|---|---|---|---|---|
| false | 3584 | true | 17737 | 189 | 13.516 | 1.03 | 1 | 1.015 |
| up | 180792 | down | 91734 | 4017 | 50.309 | 1.08 | 1 | 1.04 |
| soft | 5869 | hard | 12834 | 82 | 8.639 | 1.03 | 1.1 | 1.065 |
| right | 74117 | wrong | 15505 | 664 | 23.540 | 1.08 | 1.1 | 1.09 |
| fast | 7352 | slow | 5724 | 85 | 8.991 | 1.1 | 1.13 | 1.115 |
| top | 22744 | bottom | 7304 | 474 | 21.390 | 1.1 | 1.13 | 1.115 |
| cold | 9328 | hot | 8733 | 421 | 20.319 | 1.08 | 1.18 | 1.13 |
| new | 124227 | old | 52486 | 1022 | 21.781 | 1.08 | 1.18 | 1.13 |
| awake | 1287 | asleep | 2252 | 27 | 5.168 | 1.15 | 1.13 | 1.14 |
| last | 71886 | first | 120825 | 802 | 13.000 | 1.2 | 1.08 | 1.14 |
| exclude | 1324 | include | 15148 | 7 | 2.267 | 1.2 | 1.1 | 1.15 |
| dirty | 2666 | clean | 6281 | 30 | 5.324 | 1.18 | 1.15 | 1.165 |
| bad | 14935 | good | 81100 | 856 | 27.189 | 1.2 | 1.15 | 1.175 |
| dark | 12653 | light | 19534 | 243 | 14.796 | 1.15 | 1.2 | 1.175 |
| tall | 4329 | short | 17792 | 62 | 7.385 | 1.13 | 1.23 | 1.18 |
| inaccurate | 458 | accurate | 2887 | 5 | 2.206 | 1.28 | 1.08 | 1.18 |
| absent | 1549 | present | 14691 | 31 | 5.363 | 1.33 | 1.05 | 1.19 |
| far | 36754 | near | 17048 | 126 | 8.436 | 1.2 | 1.18 | 1.19 |
| unsuitable | 691 | suitable | 6021 | 6 | 2.364 | 1.25 | 1.15 | 1.2 |
| short | 19660 | long | 50614 | 523 | 20.902 | 1.2 | 1.23 | 1.215 |
| small | 43118 | large | 34269 | 800 | 25.674 | 1.25 | 1.18 | 1.215 |
| happy | 11340 | sad | 3322 | 33 | 5.417 | 1.23 | 1.2 | 1.215 |
| big | 24853 | small | 43118 | 210 | 10.797 | 1.1 | 1.35 | 1.225 |
| legal | 12981 | illegal | 2392 | 31 | 5.289 | 1.38 | 1.1 | 1.24 |
| little | 29018 | big | 24853 | 128 | 8.129 | 1.18 | 1.3 | 1.24 |
| out | 149187 | in | 1843054 | 1659 | -3.087 | 1.15 | 1.33 | 1.24 |
| unhappy | 1846 | happy | 11340 | 17 | 3.869 | 1.2 | 1.28 | 1.24 |
| untidy | 381 | tidy | 719 | 3 | 1.724 | 1.4 | 1.08 | 1.24 |
| rational | 2295 | irrational | 494 | 13 | 3.589 | 1.45 | 1.08 | 1.265 |
| old | 52486 | young | 32326 | 461 | 17.534 | 1.13 | 1.43 | 1.28 |
| odd | 4312 | even | 4146 | 18 | 4.104 | 1.2 | 1.4 | 1.3 |
| disorganised | 80 | organised | 767 | 0 | 0 | 1.33 | 1.28 | 1.305 |
| easy | 14414 | difficult | 21621 | 46 | 4.487 | 1.23 | 1.38 | 1.305 |
| right | 74117 | left | 13094 | 1211 | 33.406 | 1.23 | 1.38 | 1.305 |
| succeed | 2102 | fail | 3331 | 24 | 4.827 | 1.23 | 1.38 | 1.305 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| light | 19534 | heavy | 9126 | 78 | 7.823 | 1.33 | 1.28 | 1.305 |
| sick | 4333 | healthy | 3527 | 9 | 2.745 | 1.38 | 1.25 | 1.315 |
| unhealthy | 277 | healthy | 3527 | 4 | 1.975 | 1.35 | 1.3 | 1.325 |
| possible | 33656 | impossible | 6826 | 15 | 0.910 | 1.23 | 1.45 | 1.34 |
| loyal | 1330 | disloyal | 92 | 0 | 0 | 1.38 | 1.33 | 1.355 |
| over | 128305 | under | 60049 | 346 | -2.087 | 1.38 | 1.35 | 1.365 |
| less | 34026 | more | 209697 | 4123 | 58.660 | 1.5 | 1.25 | 1.375 |
| sober | 595 | drunk | 2162 | 14 | 3.724 | 1.45 | 1.3 | 1.375 |
| good | 81100 | evil | 2773 | 271 | 15.779 | 1.38 | 1.4 | 1.39 |
| interested | 8787 | uninterested | 138 | 1 | 0.9394 | 1.43 | 1.4 | 1.415 |
| misbehave | 28 | behave | 1704 | 0 | 0 | 1.43 | 1.43 | 1.43 |
| defend | 2021 | attack | 1457 | 10 | 3.068 | 1.5 | 1.38 | 1.44 |
| giant | 958 | dwarf | 572 | 3 | 1.716 | 1.55 | 1.35 | 1.45 |
| life | 54993 | death | 19891 | 551 | 21.145 | 1.5 | 1.4 | 1.45 |
| even | 4146 | uneven | 664 | 0 | 0 | 1.28 | 1.63 | 1.455 |
| take | 69607 | give | 43976 | 235 | 5.356 | 1.53 | 1.45 | 1.49 |
| tiny | 5186 | huge | 7649 | 19 | 3.904 | 1.6 | 1.38 | 1.49 |
| colossal | 235 | miniscule | 21 | 0 | 0 | 1.58 | 1.43 | 1.505 |
| pack | 871 | unpack | 114 | 0 | 0 | 1.38 | 1.65 | 1.515 |
| undress | 144 | dress | 927 | 5 | 2.23 | 1.68 | 1.35 | 1.515 |
| miniature | 858 | gigantic | 400 | 0 | 0 | 1.6 | 1.45 | 1.525 |
| hard | 15291 | easy | 14414 | 28 | 3.211 | 1.43 | 1.68 | 1.555 |
| together | 29944 | apart | 3391 | 37 | 5.248 | 1.65 | 1.55 | 1.6 |
| unorganised | 32 | organised | 3911 | 1 | 0.993 | 1.65 | 1.58 | 1.615 |
| indirectly | 1038 | directly | 8607 | 366 | 19.107 | 1.55 | 1.75 | 1.65 |
| behind | 19098 | in front | 6950 | 62 | 7.032 | 1.45 | 1.9 | 1.675 |
| buy | 12293 | sell | 7539 | 205 | 13.994 | 1.93 | 1.43 | 1.68 |
| go | 85465 | come | 66694 | 702 | 15.749 | 1.9 | 1.65 | 1.775 |
| masculine | 620 | feminine | 728 | 80 | 8.941 | 1.58 | 2 | 1.79 |
| imprecise | 120 | exact | 2152 | 0 | 0 | 1.88 | 1.8 | 1.84 |
| warm | 6744 | cool | 3822 | 36 | 5.785 | 2.08 | 1.6 | 1.84 |
| sinful | 166 | virtuous | 175 | 0 | 0 | 1.78 | 2 | 1.89 |
| married | 9779 | unmarried | 584 | 18 | 4.175 | 2.13 | 1.8 | 1.965 |
| figuratively | 45 | literally | 1936 | 5 | 2.234 | 2.23 | 1.7 | 1.965 |
| freezing | 967 | boiling | 883 | 5 | 2.216 | 2.1 | 1.85 | 1.975 |
| small | 43118 | huge | 7649 | 24 | 1.536 | 2.05 | 1.95 | 2 |
| male | 7716 | female | 4980 | 1063 | 32.544 | 1.88 | 2.18 | 2.03 |
| big | 24853 | tiny | 5186 | 13 | 1.820 | 2.28 | 1.8 | 2.04 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| large | 34269 | little | 62638 | 75 | -3.719 | 2.15 | 1.95 | 2.05 |
| interested | 8787 | disinterested | 172 | 2 | 1.360 | 2.4 | 1.75 | 2.075 |
| small | 43118 | large | 34269 | 800 | 25.674 | 3.05 | 1.15 | 2.1 |
| woman | 22008 | man | 58165 | 946 | 28.678 | 2.18 | 2.03 | 2.105 |
| excellent | 6620 | atrocious | 118 | 0 | 0 | 2.05 | 2.23 | 2.14 |
| credit | 7297 | debit | 267 | 34 | 5.814 | 2.1 | 2.25 | 2.19 |
| little | 29018 | gigantic | 400 | 1 | 0.420 | 2.1 | 2.28 | 2.19 |
| similar | 18295 | different | 47604 | 131 | 7.644 | 2.28 | 2.1 | 2.19 |
| brief | 4327 | long | 33240 | 4 | -1.846 | 2.58 | 1.83 | 2.205 |
| chilly | 336 | warm | 6082 | 2 | 1.342 | 2.45 | 2.08 | 2.265 |
| badly | 4187 | well | 108935 | 58 | 4.624 | 2.45 | 2.13 | 2.29 |
| small | 43118 | tall | 4329 | 21 | 2.548 | 1.38 | 3.23 | 2.305 |
| serious | 12093 | trivial | 873 | 10 | 2.995 | 2.45 | 2.2 | 2.325 |
| correct | 5812 | mistaken | 1099 | 0 | 0 | 2.73 | 1.95 | 2.34 |
| incorrect | 695 | right | 34396 | 0 | 0 | 2.08 | 2.6 | 2.34 |
| valley | 4613 | hill | 6881 | 71 | 3.738 | 2.55 | 2.15 | 2.35 |
| damage | 8301 | repair | 2276 | 100 | 9.905 | 2.75 | 1.98 | 2.365 |
| lend | 1254 | borrow | 1425 | 22 | 4.671 | 2.9 | 1.98 | 2.44 |
| contemporary | 4488 | ancient | 4910 | 9 | 2.633 | 2.78 | 2.15 | 2.465 |
| spicy | 207 | mild | 1543 | 0 | 0 | 2.45 | 2.53 | 2.49 |
| pale | 3237 | dark | 9777 | 47 | 6.625 | 2.58 | 2.48 | 2.53 |
| sweet | 3177 | bitter | 2353 | 20 | 4.388 | 2.48 | 2.78 | 2.63 |
| excellent | 6620 | bad | 14935 | 4 | -0.469 | 2.45 | 2.83 | 2.64 |
| work | 89441 | play | 21119 | 142 | 3.999 | 2.7 | 2.6 | 2.65 |
| hot | 8733 | cool | 3228 | 30 | 5.220 | 2.58 | 2.78 | 2.68 |
| sour | 622 | sweet | 3177 | 42 | 6.465 | 2.33 | 3.03 | 2.68 |
| single | 18074 | married | 9779 | 66 | 7.037 | 2.93 | 2.48 | 2.705 |
| ask | 18642 | answer | 14305 | 55 | 5.6201 | 2.9 | 2.55 | 2.725 |
| away | 47116 | home | 50539 | 801 | 24.099 | 2.63 | 2.88 | 2.755 |
| present | 17328 | past | 19097 | 601 | 23.841 | 2.43 | 3.13 | 2.78 |
| employer | 3002 | employee | 3108 | 109 | 10.395 | 2.93 | 2.75 | 2.84 |
| scorching | 122 | glacial | 269 | 0 | 0 | 3.35 | 2.75 | 3.05 |
| married | 9779 | divorced | 909 | 46 | 6.716 | 3 | 3.15 | 3.075 |
| mentor | 260 | protégé | 43 | 0 | 0 | 3.25 | 2.93 | 3.09 |
| pupil | 2307 | teacher | 8633 | 132 | 11.402 | 2.9 | 3.4 | 3.15 |
| aloof | 226 | amiable | 279 | 0 | 0 | 3.15 | 3.18 | 3.165 |
| wife | 16474 | husband | 10612 | 802 | 28.011 | 3.23 | 3.1 | 3.165 |
| current | 13292 | former | 16845 | 38 | 4.350 | 3.43 | 2.95 | 3.19 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| apprentice | 500 | master | 5512 | 5 | 2.174 | 3.23 | 3.2 | 3.215 |
| murderer | 768 | victim | 3796 | 10 | 3.116 | 3.33 | 3.15 | 3.24 |
| sweltering | 56 | nippy | 43 | 0 | 0 | 2.95 | 3.6 | 3.275 |
| good | 77128 | disobedient | 72 | 0 | 0 | 3.05 | 3.7 | 3.375 |
| patient | 6820 | doctor | 9048 | 102 | 9.757 | 3.5 | 3.25 | 3.375 |
| tutee | 13 | tutor | 1081 | 0 | 0 | 3.33 | 3.45 | 3.39 |
| brother | 8277 | sister | 7150 | 310 | 17.438 | 3.33 | 3.5 | 3.415 |
| gracious | 423 | mean | 2493 | 0 | 0 | 3.35 | 3.53 | 3.44 |
| flammable | 75 | inflammable | 53 | 0 | 0 | 3.18 | 3.73 | 3.455 |
| tenant | 2570 | landlord | 2673 | 338 | 18.366 | 3.25 | 3.68 | 3.465 |
| dad | 6564 | mum | 8152 | 706 | 26.470 | 3.6 | 3.4 | 3.5 |
| flee | 445 | chase | 517 | 0 | 0 | 3.68 | 3.33 | 3.505 |
| teacher | 8633 | student | 7590 | 96 | 9.463 | 3.38 | 3.73 | 3.555 |
| chilly | 336 | steaming | 393 | 0 | 0 | 3.08 | 4.05 | 3.565 |
| child | 23669 | parent | 3707 | 254 | 15.662 | 4.05 | 3.1 | 3.575 |
| mediocre | 174 | brilliant | 3411 | 0 | 0 | 3.39 | 3.83 | 3.61 |
| keep | 26723 | sell | 7539 | 19 | 2.050 | 3.88 | 3.4 | 3.64 |
| uneasy | 923 | easy | 14414 | 0 | 0 | 3.75 | 3.6 | 3.675 |
| town | 17853 | country | 31416 | 511 | 21.366 | 3.18 | 4.25 | 3.715 |
| traditional | 9714 | trendy | 239 | 0 | 0 | 4.08 | 3.4 | 3.74 |
| cow | 1326 | bull | 1808 | 5 | 2.182 | 3.98 | 3.53 | 3.755 |
| present | 18538 | future | 22378 | 206 | 12.909 | 3.48 | 4.2 | 3.84 |
| nephew | 714 | niece | 455 | 13 | 3.601 | 3.83 | 4.03 | 3.93 |
| mobile | 182 | landline | 4 | 0 | 0 | 4.58 | 3.33 | 3.955 |
| still | 4903 | sparkling | 761 | 9 | 2.937 | 3.63 | 4.33 | 3.98 |
| cold | 9328 | friendly | 3951 | 1 | -0.840 | 3.8 | 4.23 | 4.015 |
| dead | 10873 | fresh | 6614 | 4 | 0.204 | 3.85 | 4.18 | 4.015 |
| donate | 230 | steal | 869 | 0 | 0 | 4.18 | 4.1 | 4.14 |
| suitable | 6021 | wrong | 15505 | 0 | 0 | 4.33 | 4 | 4.165 |
| helpful | 3115 | helpless | 792 | 0 | 0 | 4.45 | 4.15 | 4.3 |
| fair | 8373 | poor | 14563 | 9 | 0.969 | 4.63 | 4.03 | 4.33 |
| double | 7070 | single | 18074 | 146 | 11.554 | 4.25 | 4.43 | 4.34 |
| white | 23427 | coloured | 2433 | 43 | 6.123 | 4.48 | 4.2 | 4.34 |
| actress | 1046 | actor | 2003 | 9 | 2.965 | 4.3 | 4.5 | 4.4 |
| bad | 14935 | satisfactory | 2161 | 0 | 0 | 4.58 | 4.25 | 4.415 |
| finish | 2824 | continue | 11641 | 3 | 0.784 | 4.23 | 4.75 | 4.49 |
| uncle | 3350 | aunt | 2744 | 104 | 10.153 | 4.65 | 4.4 | 4.525 |
| father | 22744 | mother | 24201 | 954 | 29.996 | 4.48 | 4.63 | 4.555 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| wintry | 124 | sultry | 119 | 0 | 0 | 4.88 | 4.25 | 4.565 |
| helpless | 792 | helpful | 3115 | 0 | 0 | 4.83 | 4.43 | 4.63 |
| escape | 4613 | hunt | 2601 | 0 | 0 | 4.73 | 4.8 | 4.765 |
| king | 15765 | queen | 7717 | 244 | 15.23 | 4.8 | 4.75 | 4.775 |
| cold | 9328 | mild | 1543 | 3 | 1.316 | 4.58 | 5.05 | 4.815 |
| good | 77128 | mediocre | 174 | 8 | 2.591 | 4.63 | 5 | 4.815 |
| bland | 608 | hot | 8733 | 0 | 0 | 4.58 | 5.13 | 4.855 |
| dry | 5273 | sweet | 3177 | 11 | 3.064 | 4.58 | 5.15 | 4.8675 |
| tomorrow | 8893 | today | 25855 | 245 | 14.918 | 4.93 | 4.83 | 4.88 |
| occupier | 382 | owner | 4957 | 21 | 4.561 | 5.08 | 4.78 | 4.93 |
| dog | 7814 | cat | 3847 | 113 | 10.488 | 4.73 | 5.38 | 5.055 |
| red | 14569 | green | 14199 | 435 | 20.361 | 5.25 | 4.88 | 5.065 |
| daughter | 9171 | mother | 24202 | 257 | 15.339 | 5.13 | 5.03 | 5.08 |
| next to | 2849 | opposite | 4890 | 4 | 1.652 | 5.2 | 5.18 | 5.19 |
| trade | 19841 | purchase | 4433 | 7 | 0.985 | 5.13 | 5.48 | 5.305 |
| short | 17792 | polite | 1092 | 1 | 0.029 | 4.9 | 5.8 | 5.35 |
| father | 22557 | daughter | 9171 | 119 | 9.961 | 5.28 | 5.55 | 5.415 |
| hungry | 1786 | thirsty | 271 | 18 | 4.236 | 5.4 | 5.43 | 5.415 |
| blue | 10059 | orange | 1790 | 28 | 5.121 | 5.45 | 5.65 | 5.55 |
| buy | 12293 | spend | 7323 | 10 | 1.740 | 5.15 | 6.05 | 5.6 |
| rent | 3440 | let | 24026 | 18 | 3.269 | 5.3 | 5.93 | 5.615 |
| walk | 10046 | crawl | 403 | 5 | 2.145 | 5.78 | 5.55 | 5.665 |
| lease | 2204 | let | 24026 | 6 | 1.369 | 5.4 | 6.25 | 5.825 |
| answer | 14305 | reply | 4225 | 3 | -0.010 | 5.55 | 6.13 | 5.84 |
| yellow | 4366 | blue | 10059 | 204 | 14.129 | 6 | 5.73 | 5.865 |
| egg | 2436 | chicken | 2042 | 51 | 7.106 | 5.95 | 5.8 | 5.875 |
| fish | 10222 | fowl | 139 | 13 | 3.585 | 5.88 | 5.9 | 5.89 |
| coffee | 6286 | tea | 8030 | 467 | 21.493 | 5.7 | 6.1 | 5.9 |
| similar | 18295 | identical | 2148 | 66 | 7.882 | 6 | 6.15 | 6.075 |
| brick | 1803 | mortar | 603 | 5 | 2.211 | 6.08 | 6.15 | 6.115 |
| clean | 6281 | tidy | 719 | 78 | 8.806 | 6.03 | 6.2 | 6.115 |
| lukewarm | 169 | tepid | 81 | 1 | 0.999 | 5.88 | 6.55 | 6.215 |
| long | 33258 | narrow | 4711 | 175 | 12.637 | 6.2 | 6.35 | 6.275 |
| glove | 375 | hand | 32575 | 57 | 7.469 | 6.33 | 6.23 | 6.28 |
| square | 4012 | long | 50614 | 11 | 0.258 | 6.13 | 6.45 | 6.29 |
| book | 23916 | page | 10709 | 98 | 8.607 | 6.3 | 6.3 | 6.3 |
| mug | 685 | cup | 11913 | 6 | 2.283 | 6.3 | 6.33 | 6.315 |
| yell | 215 | call | 18778 | 0 | 0 | 6.1 | 6.55 | 6.325 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| rain | 6253 | wind | 7357 | 217 | 14.574 | 6.45 | 6.28 | 6.365 |
| tipsy | 65 | inebriated | 31 | 0 | 0 | 6.38 | 6.35 | 6.365 |
| carrot | 377 | pea | 173 | 1 | 0.996 | 6.55 | 6.2 | 6.375 |
| finger | 3044 | hand | 32576 | 94 | 9.184 | 6.38 | 6.43 | 6.405 |
| humid | 208 | hot | 8733 | 35 | 5.900 | 6.4 | 6.63 | 6.515 |
| soft | 5869 | easy | 14414 | 18 | 3.246 | 6.33 | 6.7 | 6.515 |
| fit | 2942 | hot | 8733 | 0 | 0 | 6.5 | 6.55 | 6.525 |
| short | 19660 | squat | 273 | 8 | 2.733 | 6.3 | 6.78 | 6.54 |
| cold | 9328 | distant | 2772 | 23 | 4.526 | 6.38 | 6.75 | 6.565 |
| cold | 9328 | sticky | 809 | 1 | 0.623 | 6.6 | 6.6 | 6.6 |
| wrong | 15505 | bad | 14935 | 21 | 2.058 | 6.5 | 6.7 | 6.6 |
| scream | 952 | kick | 2269 | 5 | 2.187 | 6.5 | 6.73 | 6.615 |
| quiet | 5482 | still | 2323 | 9 | 2.721 | 6.68 | 6.58 | 6.63 |
| tired | 3852 | sick | 4333 | 68 | 8.145 | 6.48 | 6.78 | 6.63 |
| dead | 10873 | gone | 18474 | 53 | 5.901 | 6.55 | 6.73 | 6.64 |
| mad | 2966 | angry | 4015 | 2 | 0.993 | 6.58 | 6.7 | 6.64 |
| sad | 3322 | unhappy | 1846 | 8 | 2.720 | 6.55 | 6.73 | 6.64 |
| crown | 5210 | king | 15765 | 40 | 5.675 | 6.75 | 6.73 | 6.74 |
| rose | 4963 | flower | 2126 | 13 | 3.459 | 6.85 | 6.65 | 6.75 |
| animal | 6611 | dog | 7814 | 21 | 4.0195 | 6.73 | 6.9 | 6.815 |
| clever | 2237 | bright | 5278 | 5 | 1.972 | 6.83 | 6.8 | 6.815 |
| easy | 14414 | uneasy | 923 | 0 | 0 | 3.33 | 3.2 | 3.265 |

**Appendix 4** Target and control stimuli for Experiment 1

| TEST | L | PH | S | IMG | FRQ | STR | CONTROL | L | PH | S | IMG | FRQ | STR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | |
| soft | 4 | 4 | 1 | 476 | 61 | 0 | bright | 6 | 4 | 1 | 489 | 87 | 0 |
| hard | 4 | 3 | 1 | 460 | 202 | 0 | clear | 5 | 3 | 1 | 456 | 219 | 0 |
| | | | | | | | | | | | | | |
| cold | 4 | 4 | 1 | 531 | 171 | 0 | fresh | 5 | 4 | 1 | 453 | 82 | 0 |
| hot | 3 | 3 | 1 | 551 | 130 | 0 | firm | 4 | 3 | 1 | 404 | 109 | 0 |
| | | | | | | | | | | | | | |
| new | 3 | 3 | 1 | 418 | 1635 | 0 | well | 4 | 3 | 1 | 522 | 897 | 0 |
| old | 3 | 3 | 1 | 478 | 660 | 0 | safe | 4 | 3 | 1 | 474 | 58 | 0 |
| | | | | | | | | | | | | | |
| right | 5 | 3 | 1 | 372 | 613 | 0 | nice | 4 | 3 | 1 | 375 | 75 | 0 |
| wrong | 5 | 3 | 1 | 344 | 129 | 0 | keen | 4 | 3 | 1 | 335 | 11 | 0 |
| | | | | | | | | | | | | | |
| big | 3 | 3 | 1 | 463 | 360 | 0 | wide | 4 | 3 | 1 | 455 | 125 | 0 |
| small | 5 | 4 | 1 | 447 | 542 | 0 | broad | 5 | 4 | 1 | 463 | 84 | 0 |
| | | | | | | | | | | | | | |
| old | 3 | 3 | 1 | 478 | 660 | 0 | safe | 4 | 3 | 1 | 474 | 58 | 0 |
| young | 5 | 3 | 1 | 521 | 385 | 0 | dead | 4 | 3 | 1 | 520 | 174 | 0 |
| | | | | | | | | | | | | | |
| tall | 4 | 3 | 1 | 514 | 55 | 0 | pale | 4 | 3 | 1 | 532 | 58 | 0 |
| short | 5 | 3 | 1 | 431 | 212 | 0 | dark | 4 | 3 | 1 | 586 | 185 | 0 |
| | | | | | | | | | | | | | |
| little | 5 | 4 | 2 | 502 | 831 | 20 | heavy | 5 | 4 | 2 | 495 | 110 | 20 |
| big | 3 | 3 | 1 | 463 | 360 | 0 | wide | 4 | 3 | 1 | 455 | 125 | 0 |
| | | | | | | | | | | | | | |
| short | 5 | 3 | 1 | 431 | 212 | 0 | dark | 4 | 3 | 1 | 586 | 185 | 0 |
| long | 4 | 3 | 1 | 471 | 755 | 0 | light | 5 | 3 | 1 | 542 | 333 | 0 |
| | | | | | | | | | | | | | |
| long | 4 | 3 | 1 | 471 | 755 | 0 | light | 5 | 3 | 1 | 542 | 333 | 0 |
| tall | 4 | 3 | 1 | 514 | 55 | 0 | pale | 4 | 3 | 1 | 532 | 58 | 0 |
| | | | | | | | | | | | | | |
| unhappy | 7 | 6 | 3 | x | 26 | O20 | untidy | 6 | 6 | 3 | 458 | 1 | 200 |
| happy | 5 | 4 | 2 | 511 | 98 | 20 | muddy | 5 | 4 | 2 | 522 | 10 | 20 |
| | | | | | | | | | | | | | |
| happy | 5 | 4 | 2 | 511 | 98 | 20 | muddy | 5 | 4 | 2 | 522 | 10 | 20 |
| sad | 3 | 3 | 1 | 419 | 35 | 0 | calm | 4 | 3 | 1 | 439 | 35 | 0 |
| | | | | | | | | | | | | | |
| odd | 3 | 2 | 1 | X | 44 | 0 | ill | 3 | 2 | 1 | X | 39 | 0 |
| even | 4 | 4 | 2 | 338 | 1171 | 20 | ideal | 5 | 4 | 2 | 331 | 61 | 20 |
| | | | | | | | | | | | | | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| bad | 3 | 3 | 1 | 388 | 142 | 0 | late | 4 | 3 | 1 | 387 | 179 | 0 |
| good | 4 | 3 | 1 | 374 | 807 | 0 | free | 4 | 3 | 1 | 397 | 260 | 0 |
| | | | | | | | | | | | | | |
| right | 5 | 3 | 1 | 372 | 613 | 0 | nice | 4 | 3 | 1 | 375 | 75 | 0 |
| left | 4 | 4 | 1 | 383 | 480 | 0 | kind | 4 | 4 | 1 | 383 | 313 | 0 |
| | | | | | | | | | | | | | |
| small | 5 | 4 | 1 | 447 | 542 | 0 | broad | 5 | 4 | 1 | 463 | 84 | 0 |
| large | 5 | 3 | 1 | 449 | 361 | 0 | full | 4 | 3 | 1 | 437 | 230 | 0 |
| | | | | | | | | | | | | | |
| even | 4 | 4 | 2 | 338 | 1171 | 20 | ideal | 5 | 4 | 2 | 331 | 61 | 20 |
| uneven | 6 | 6 | 3 | X | 6 | 220 | immature | 8 | 6 | 3 | 439 | 3 | 102 |
| | | | | | | | | | | | | | |
| disorganised | 12 | 10 | 4 | X | X | 2200 | conservative | 12 | 10 | 4 | 373 | 31 | O200 |
| organised | 9 | 7 | 3 | X | X | 200 | inferior | 8 | 7 | 3 | 379 | 7 | O20 |
| | | | | | | | | | | | | | |
| organised | 9 | 7 | 3 | X | X | 200 | inferior | 8 | 7 | 3 | 379 | 7 | O20 |
| unorganised | 11 | 9 | 4 | X | X | 2200 | unpopular | 9 | 9 | 4 | 307 | 6 | O200 |
| | | | | | | | | | | | | | |
| unhealthy | 9 | 7 | 3 | X | 4 | O20 | unstable | 8 | 7 | 3 | 356 | 8 | 220 |
| healthy | 7 | 5 | 2 | X | 33 | 20 | angry | 5 | 5 | 2 | 492 | 45 | 20 |
| | | | | | | | | | | | | | |
| pack | 4 | 3 | 1 | X | 25 | 0 | spare | 5 | 3 | 1 | 316 | 8 | 0 |
| unpack | 6 | 5 | 2 | X | 1 | 22 | offend | 6 | 5 | 2 | 372 | 4 | O2 |
| | | | | | | | | | | | | | |
| sick | 4 | 3 | 1 | 456 | 51 | 0 | mad | 3 | 3 | 1 | 479 | 39 | 0 |
| healthy | 7 | 5 | 2 | X | 33 | 20 | angry | 5 | 5 | 2 | 492 | 45 | 20 |
| | | | | | | | | | | | | | |
| good | 4 | 3 | 1 | 374 | 807 | 0 | free | 4 | 3 | 1 | 397 | 260 | 0 |
| evil | 4 | 3 | 2 | 434 | 72 | 20 | idle | 4 | 3 | 2 | 444 | 13 | 20 |
| | | | | | | | | | | | | | |
| hard | 4 | 3 | 1 | 460 | 202 | 0 | clear | 5 | 3 | 1 | 456 | 219 | 0 |
| easy | 4 | 3 | 2 | 321 | 125 | 20 | eager | 5 | 3 | 2 | 401 | 27 | 20 |
| | | | | | | | | | | | | | |
| interested | 10 | 9 | 3 | X | 105 | 200 | competent | 9 | 9 | 3 | 295 | 21 | 200 |
| uninterested | 12 | 11 | 4 | X | 1 | 2200 | incompetent | 11 | 11 | 4 | 359 | 2 | 2000 |
| | | | | | | | | | | | | | |
| disinterested | 13 | 12 | 4 | X | 5 | O200 | inconsistent | 12 | 12 | 4 | X | 5 | 1020 |
| interested | 10 | 9 | 3 | X | 105 | 200 | competent | 9 | 9 | 3 | 295 | 21 | 200 |
| | | | | | | | | | | | | | |
| defend | 6 | 6 | 2 | X | 21 | O2 | insult | 6 | 6 | 2 | 477 | 7 | 20 |
| attack | 6 | 4 | 2 | 501 | 105 | O2 | repair | 6 | 4 | 2 | 440 | 20 | O2 |
| | | | | | | | | | | | | | |
| less | 4 | 3 | 1 | 286 | 438 | 0 | soon | 4 | 3 | 1 | 278 | 199 | 0 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| more | 4 | 2 | 1 | 264 | 340 | 0 | here | 4 | 2 | 1 | 278 | 750 | 0 |
| | | | | | | | | | | | | | |
| take | 4 | 3 | 1 | 337 | 611 | 0 | put | 3 | 3 | 1 | 263 | 437 | 0 |
| give | 4 | 3 | 1 | 383 | 391 | 0 | need | 4 | 3 | 1 | 327 | 360 | 0 |
| | | | | | | | | | | | | | |
| giant | 5 | 5 | 2 | 562 | 23 | 20 | village | 7 | 5 | 2 | 578 | 72 | 20 |
| dwarf | 5 | 4 | 1 | X | 3 | 0 | tribe | 5 | 4 | 1 | 515 | 12 | 0 |
| | | | | | | | | | | | | | |
| colossal | 8 | 6 | 3 | X | 3 | O20 | fidgety | 7 | 6 | 3 | 407 | X | 200 |
| miniscule | 9 | 9 | 3 | X | X | 200 | practical | 9 | 9 | 3 | 358 | 68 | 200 |
| | | | | | | | | | | | | | |
| masculine | 9 | 9 | 3 | X | 7 | 200 | impatient | 9 | 8 | 3 | X | 10 | O20 |
| feminine | 8 | 7 | 3 | X | 10 | 200 | neurotic | 8 | 8 | 3 | 417 | 10 | O20 |
| | | | | | | | | | | | | | |
| male | 4 | 3 | 1 | 587 | 37 | 0 | gray | 4 | 3 | 1 | 541 | 80 | O20 |
| female | 6 | 5 | 2 | | 50 | 20 | modern | 6 | 5 | 2 | 368 | 198 | 20 |
| | | | | | | | | | | | | | |
| go | 2 | 2 | 1 | 364 | 626 | 0 | know | 4 | 2 | 1 | 278 | 683 | 0 |
| come | 4 | 3 | 1 | 322 | 630 | 0 | guess | 5 | 3 | 1 | 330 | 56 | 0 |
| | | | | | | | | | | | | | |
| buy | 3 | 2 | 1 | 397 | 70 | 0 | see | 3 | 2 | 1 | 379 | 772 | 0 |
| sell | 4 | 3 | 1 | 397 | 41 | 0 | look | 4 | 3 | 1 | 395 | 399 | 0 |
| | | | | | | | | | | | | | |
| warm | 4 | 3 | 1 | 441 | 67 | 0 | thick | 5 | 3 | 1 | 468 | 67 | 0 |
| cool | 4 | 3 | 1 | 429 | 62 | 0 | thin | 4 | 3 | 1 | 502 | 92 | 0 |
| | | | | | | | | | | | | | |
| freezing | 8 | 6 | 2 | X | 15 | 20 | culture | 7 | 6 | 2 | 339 | 58 | 20 |
| boiling | 7 | 5 | 2 | X | 9 | 20 | foreign | 7 | 5 | 2 | 404 | 158 | 20 |
| | | | | | | | | | | | | | |
| credit | 6 | 6 | 2 | X | 64 | 20 | talent | 6 | 6 | 2 | 399 | 40 | 20 |
| debit | 5 | 5 | 2 | X | X | 20 | belief | 6 | 5 | 2 | 328 | 64 | O2 |
| | | | | | | | | | | | | | |
| large | 5 | 3 | 1 | 449 | 361 | 0 | full | 4 | 3 | 1 | 437 | 230 | 0 |
| little | 5 | 4 | 2 | 502 | 831 | 20 | heavy | 5 | 4 | 2 | 495 | 110 | 20 |
| | | | | | | | | | | | | | |
| woman | 5 | 5 | 2 | 626 | 224 | 20 | winter | 6 | 5 | 2 | 621 | 83 | 20 |
| man | 3 | 3 | 1 | 567 | 1207 | 0 | south | 5 | 3 | 1 | 476 | 240 | 0 |
| | | | | | | | | | | | | | |
| big | 3 | 3 | 1 | 463 | 360 | 0 | wide | 4 | 3 | 1 | 455 | 125 | 0 |
| tiny | 4 | 4 | 2 | X | 50 | 20 | ugly | 4 | 4 | 2 | 491 | 21 | 20 |
| | | | | | | | | | | | | | |
| work | 4 | 3 | 1 | 458 | 760 | 0 | life | 4 | 3 | 1 | 482 | 715 | 0 |
| play | 4 | 3 | 1 | 498 | 200 | 0 | line (note) | 4 | 3 | 1 | 491 | 298 | 0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| pupil | 5 | 5 | 2 | 572 | 20 | 20 | engine | 6 | 5 | 2 | 595 | 50 | 20 |
| teacher | 7 | 5 | 2 | 575 | 80 | 20 | machine | 7 | 5 | 2 | 575 | 103 | O2 |
| | | | | | | | | | | | | | |
| employer | 8 | 6 | 3 | X | 15 | O20 | receiver | 8 | 6 | 3 | 504 | 13 | O20 |
| employee | 8 | 6 | 3 | X | 24 | 102 | engineer | 8 | 6 | 3 | 495 | 42 | 102 |
| | | | | | | | | | | | | | |
| wife | 4 | 3 | 1 | 575 | 228 | 0 | town | 4 | 3 | 1 | 553 | 212 | 0 |
| husband | 7 | 7 | 2 | 537 | 131 | 20 | subject | 7 | 7 | 2 | 418 | 161 | 20 |
| | | | | | | | | | | | | | |
| tenant | 6 | 6 | 2 | X | 5 | 20 | critic | 6 | 6 | 2 | X | 25 | 20 |
| landlord | 8 | 6 | 2 | X | 12 | 20 | fishwife | 8 | 6 | 2 | X | X | 20 |
| | | | | | | | | | | | | | |
| brother | 7 | 5 | 2 | 589 | 73 | 20 | missile | 7 | 5 | 2 | 602 | 48 | 20 |
| sister | 6 | 5 | 2 | 613 | 38 | 20 | bullet | 6 | 5 | 2 | 611 | 28 | 20 |
| | | | | | | | | | | | | | |
| teacher | 7 | 5 | 2 | 575 | 80 | 20 | machine | 7 | 5 | 2 | 575 | 103 | O2 |
| student | 7 | 8 | 2 | 603 | 131 | 20 | context | 7 | 8 | 2 | 274 | 35 | 20 |
| | | | | | | | | | | | | | |
| patient | 7 | 6 | 2 | 526 | 86 | 20 | captain | 7 | 6 | 2 | 497 | 85 | 20 |
| doctor | 6 | 5 | 2 | 600 | 100 | 20 | driver | 6 | 5 | 2 | 567 | 49 | 20 |
| | | | | | | | | | | | | | |
| dad | 3 | 3 | 1 | 626 | 15 | 0 | pig | 3 | 3 | 1 | 635 | 8 | 0 |
| mum | 3 | 3 | 1 | | 1 | 0 | hog | 3 | 3 | 1 | 527 | 3 | 0 |
| | | | | | | | | | | | | | |
| flee | 4 | 3 | 1 | 431 | 1 | 0 | zone | 4 | 3 | 1 | 432 | 11 | 0 |
| chase | 5 | 4 | 1 | X | 18 | 0 | steam | 5 | 4 | 1 | 591 | 17 | 0 |
| | | | | | | | | | | | | | |
| cow | 3 | 2 | 1 | 632 | 29 | 0 | ice | 3 | 2 | 1 | 635 | 45 | 0 |
| bull | 4 | 3 | 1 | X | 14 | 0 | snow | 4 | 3 | 1 | 597 | 59 | 0 |
| | | | | | | | | | | | | | |
| child | 5 | 5 | 1 | 619 | 213 | 0 | plant | 5 | 5 | 1 | 605 | 125 | 0 |
| parent | 6 | 6 | 2 | X | 15 | 20 | forest | 6 | 6 | 2 | 633 | 66 | 20 |
| | | | | | | | | | | | | | |
| actress | 7 | 6 | 2 | X | 6 | 20 | servant | 7 | 6 | 2 | 508 | 19 | 20 |
| actor | 5 | 4 | 2 | X | 24 | 20 | widow | 5 | 4 | 2 | 505 | 26 | 20 |
| | | | | | | | | | | | | | |
| father | 6 | 4 | 2 | 646 | 183 | 20 | dollar | 6 | 4 | 2 | 611 | 46 | 20 |
| mother | 6 | 4 | 2 | 638 | 216 | 20 | penny | 5 | 4 | 2 | 609 | 25 | 20 |
| | | | | | | | | | | | | | |
| cold | 4 | 4 | 1 | 531 | 171 | 0 | fresh | 5 | 4 | 1 | 453 | 82 | 0 |
| mild | 4 | 4 | 1 | 359 | 14 | 0 | clean | 5 | 4 | 1 | 454 | 70 | 0 |
| | | | | | | | | | | | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| king | 4 | 3 | 1 | 585 | 88 | 0 | song | 4 | 3 | 1 | 578 | 70 | 0 |
| queen | 5 | 4 | 1 | 612 | 41 | 0 | clock | 5 | 4 | 1 | 614 | 20 | 0 |
| | | | | | | | | | | | | | |
| occupier | 8 | 7 | 4 | X | X | 2000 | mediator | 8 | 7 | 4 | X | X | 2000 |
| owner | 5 | 3 | 2 | 425 | 33 | 20 | order | 5 | 3 | 2 | 352 | 376 | 20 |
| | | | | | | | | | | | | | |
| hungry | 6 | 6 | 2 | 503 | 23 | 20 | mystic | 6 | 6 | 2 | 495 | 3 | 20 |
| thirsty | 7 | 5 | 2 | 482 | 5 | 20 | lovely | 6 | 5 | 2 | 491 | 44 | 20 |
| | | | | | | | | | | | | | |
| coffee | 6 | 4 | 2 | 618 | 78 | 20 | ocean | 5 | 4 | 2 | 623 | 34 | 20 |
| tea | 3 | 2 | 1 | 599 | 28 | 0 | sea | 3 | 2 | 1 | 606 | 95 | 0 |
| | | | | | | | | | | | | | |
| married | 7 | 5 | 2 | 568 | 105 | 20 | empty | 5 | 5 | 2 | 479 | 64 | 20 |
| single | 6 | 5 | 2 | 415 | 172 | 20 | quiet | 5 | 5 | 2 | 426 | 76 | 20 |
| | | | | | | | | | | | | | |
| unmarried | 9 | 7 | 3 | X | 6 | 220 | impartial | 9 | 7 | 3 | 305 | 8 | 220 |
| married | 7 | 5 | 2 | 568 | 105 | 20 | empty | 5 | 5 | 2 | 479 | 64 | 20 |
| | | | | | | | | | | | | | |
| sweet | 5 | 4 | 1 | 493 | 70 | 0 | wild | 4 | 4 | 1 | 500 | 56 | 0 |
| bitter | 6 | 4 | 2 | 457 | 53 | 20 | funny | 5 | 4 | 2 | 468 | 41 | 20 |
| | | | | | | | | | | | | | |
| sour | 4 | 3 | 2 | 495 | 3 | 20 | eerie | 5 | 3 | 2 | 423 | 2 | 20 |
| sweet | 5 | 4 | 1 | 493 | 70 | 0 | wild | 4 | 4 | 1 | 500 | 56 | 0 |
| | | | | | | | | | | | | | |

| | | | |
|---|---|---|---|
| **L** | number of letters | **IMG** | imageability |
| **PH** | number of phonemes | **FRQ** | frequency (Francis & Kucera 1982) |
| **S** | number of syllables | **STR** | stress pattern |

**Appendix 5** Sample instructions for Experiment 1

## LEXICAL OPPOSITION – LEXICAL DECISION TASK I

Sandra Kotzor
Kellogg College, University of Oxford
sandra.kotzor@ling-phil.ox.ac.uk

# Instructions

***Please read the following instructions carefully before the task.***

In the task you are about to participate in you will be shown pairs of words. Your task is to decide whether these words are opposites or not. If you DO consider them to be opposites (use your own intuition), you press the **YES** key. If you DO NOT consider them to be opposites you press the **NO** key. Please use the thumbs of both hands to press the buttons (like on a games console).

**Example:**

| | |
|---|---|
| MAXIMIZE - MINIMIZE | YES |
| UNHAPPY - SAD | NO |

**Procedure:**

First of all please make sure that you hand your consent form to the experimenter if she has not already taken it. When you sit down, please ensure you are able to see the screen well and without reflection.

First there will be a trial task of 10 items to ensure the all instructions are clear. This task will be repeated TWICE with the same words pairs. After the trial task you will have a last chance to ask the experimenter any questions you might have. The experimenter will remain in the room with you during the main task (in case of technical problems) but will not interfere.

The main task consists of 140 items and the task will take approximately 7 minutes. The two words will appear next to each other on the screen at the same time (as in the examples above). Please make your decision as quickly as possible. Some of the words will appear several times – don't worry if you think you have seen them before. However, none of the pairs are the same.

This is not a test, so please relax and don't worry about doing it 'right' or 'wrong'. There is no right and wrong in this task.

Thank you very much for your participation – I hope you enjoy the task.

**Appendix 6** Target and control stimuli for Experiment 2

| TEST | L | PH | S | IMG | FRQ | STR | CONTROL | L | PH | S | IMG | FRQ | STR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | |
| happy | 5 | 4 | 2 | 511 | 98 | 20 | muddy | 5 | 4 | 2 | 522 | 10 | 20 |
| sad | 3 | 3 | 1 | 419 | 35 | 0 | calm | 4 | 3 | 1 | 439 | 35 | 0 |
| unhappy | 7 | 6 | 3 | x | 26 | O20 | untidy | 6 | 6 | 3 | 458 | 1 | 200 |
| | | | | | | | | | | | | | |
| healthy | 7 | 5 | 2 | X | 33 | 20 | angry | 5 | 5 | 2 | 492 | 45 | 20 |
| unhealthy | 9 | 7 | 3 | X | 4 | O20 | unstable | 8 | 7 | 3 | 356 | 8 | 220 |
| sick | 4 | 3 | 1 | 456 | 51 | 0 | mad | 3 | 3 | 1 | 479 | 39 | 0 |
| | | | | | | | | | | | | | |
| organised | 9 | 7 | 3 | X | X | 200 | inferior | 8 | 7 | 3 | 379 | 7 | O20 |
| unorganised | 11 | 9 | 4 | X | X | 2200 | unpopular | 9 | 9 | 4 | 307 | 6 | O200 |
| disorganised | 12 | 10 | 4 | X | X | 2200 | conservative | 12 | 10 | 4 | 373 | 31 | O200 |
| | | | | | | | | | | | | | |
| interested | 10 | 9 | 3 | X | 105 | 200 | competent | 9 | 9 | 3 | 295 | 21 | 200 |
| uninterested | 12 | 11 | 4 | X | 1 | 2200 | incompetent | 11 | 11 | 4 | 359 | 2 | 2000 |
| disinterested | 13 | 12 | 4 | X | 5 | O200 | inconsistent | 12 | 12 | 4 | X | 5 | 1020 |
| | | | | | | | | | | | | | |
| short | 5 | 3 | 1 | 431 | 212 | 0 | dark | 4 | 3 | 1 | 586 | 185 | 0 |
| tall | 4 | 3 | 1 | 514 | 55 | 0 | pale | 4 | 3 | 1 | 532 | 58 | 0 |
| long | 4 | 3 | 1 | 471 | 755 | 0 | light | 5 | 3 | 1 | 542 | 333 | 0 |
| | | | | | | | | | | | | | |
| little | 5 | 4 | 2 | 502 | 831 | 20 | heavy | 5 | 4 | 2 | 495 | 110 | 20 |
| big | 3 | 3 | 1 | 463 | 360 | 0 | wide | 4 | 3 | 1 | 455 | 125 | 0 |
| small | 5 | 4 | 1 | 447 | 542 | 0 | broad | 5 | 4 | 1 | 463 | 84 | 0 |
| large | 5 | 3 | 1 | 449 | 361 | 0 | full | 4 | 3 | 1 | 437 | 230 | 0 |
| | | | | | | | | | | | | | |
| old | 3 | 3 | 1 | 478 | 660 | 0 | safe | 4 | 3 | 1 | 474 | 58 | 0 |
| new | 3 | 3 | 1 | 418 | 1635 | 0 | well | 4 | 3 | 1 | 522 | 897 | 0 |
| young | 5 | 3 | 1 | 521 | 385 | 0 | dead | 4 | 3 | 1 | 520 | 174 | 0 |
| | | | | | | | | | | | | | |
| even | 4 | 4 | 2 | 338 | 1171 | 20 | ideal | 5 | 4 | 2 | 331 | 61 | 20 |
| odd | 3 | 2 | 1 | X | 44 | 0 | ill | 3 | 2 | 1 | X | 39 | 0 |
| uneven | 6 | 6 | 3 | X | 6 | 220 | immature | 8 | 6 | 3 | 439 | 3 | 102 |
| | | | | | | | | | | | | | |
| married | 7 | 5 | 2 | 568 | 105 | 20 | empty | 5 | 5 | 2 | 479 | 64 | 20 |
| unmarried | 9 | 7 | 3 | X | 6 | 220 | impartial | 9 | 7 | 3 | 305 | 8 | 220 |
| single | 6 | 5 | 2 | 415 | 172 | 20 | quiet | 5 | 5 | 2 | 426 | 76 | 20 |
| | | | | | | | | | | | | | |
| sweet | 5 | 4 | 1 | 493 | 70 | 0 | wild | 4 | 4 | 1 | 500 | 56 | 0 |
| bitter | 6 | 4 | 2 | 457 | 53 | 20 | funny | 5 | 4 | 2 | 468 | 41 | 20 |
| sour | 4 | 3 | 2 | 495 | 3 | 20 | eerie | 5 | 3 | 2 | 423 | 2 | 20 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | |
| hard | 4 | 3 | 1 | 460 | 202 | 0 | clear | 5 | 3 | 1 | 456 | 219 | 0 |
| easy | 4 | 3 | 2 | 321 | 125 | 20 | eager | 5 | 3 | 2 | 401 | 27 | 20 |
| soft | 4 | 4 | 1 | 476 | 61 | 0 | bright | 6 | 4 | 1 | 489 | 87 | 0 |
| | | | | | | | | | | | | |
| good | 4 | 3 | 1 | 374 | 807 | 0 | free | 4 | 3 | 1 | 397 | 260 | 0 |
| bad | 3 | 3 | 1 | 388 | 142 | 0 | late | 4 | 3 | 1 | 387 | 179 | 0 |
| evil | 4 | 3 | 2 | 434 | 72 | 20 | idle | 4 | 3 | 2 | 444 | 13 | 20 |
| | | | | | | | | | | | | |
| right | 5 | 3 | 1 | 372 | 613 | 0 | nice | 4 | 3 | 1 | 375 | 75 | 0 |
| left | 4 | 4 | 1 | 383 | 480 | 0 | kind | 4 | 4 | 1 | 383 | 313 | 0 |
| wrong | 5 | 3 | 1 | 344 | 129 | 0 | keen | 4 | 3 | 1 | 335 | 11 | 0 |
| | | | | | | | | | | | | |

| | | | |
|---|---|---|---|
| **L** | number of letters | **IMG** | imageability |
| **PH** | number of phonemes | **FRQ** | frequency (Francis & Kucera 1982) |
| **S** | number of syllables | **STR** | stress pattern |

## **Appendix 7** Sample of the attribute listing task

Please read the following instructions carefully before starting the questionnaire.

Which attributes or characteristics do you think describe each of the following objects or concepts?

Please list all attributes or characteristics you think are relevant in the box below the corresponding word in the order they come to mind (like in the example of ROBIN below).

Don't worry if you have longer lists for some words than for others – that is completely normal. Please try to spend no longer than one minute on each item.

**EXAMPLE: ROBIN**
is small, lays eggs, has wings, has two legs, has feathers, flies, chirps/sings, has a beak, is delicate

Thank you very much for your help with this questionnaire!

**1. Please list as many characteristics as you can think of in about one minute for each of the objects or concepts below.**

| TASK A | TASK B |
|---|---|
| CARROT | PEA |
| COW | BULL |
| APPRENTICE | MASTER |
| COFFEE | TEA |
| KING | QUEEN |
| DOG | CAT |
| EMPLOYER | EMPLOYEE |
| GLOVE | HAND |
| FATHER | MOTHER |
| MUG | CUP |
| GIANT | DWARF |
| RAIN | WIND |
| WOMAN | MAN |
| VALLEY | HILL |
| STUDENT | TEACHER |

## **Appendix 8** English attribute listing results

**PEA**

green, round, uncomfortable for princesses, tastes bad, small
small, round, green, squishy, tasty, healthy, vegetable
small, green, round, tasty , seed
green, round, kept princess awake, healthy, frozen, comes in pods
is green, tastes nice, is good for you, is a vegetable, in a pod
is green, is round, tastes yummy, in a pod, rolls
green, round, squishy, associated with princesses, good for you, tastes good, goes with fish, grows in the garden in pod
green, comes in pod, uniform, round, small, likes hanging out in groups, rolls, has dimples, can be dried, mushy, sweet
green, small, round, comes in a pod
comes in pod, is green, edible, tastes sweet, grows on a vine
plant has flowers, green, small, comes in a pod, sweet
tiny, pearlish, has company, is green, can be smashed, is protected

**CARROT**

orange, earth, long feathery leaves, knobbly, healthy
orange, long, root vegetable, nutritious
orange with green feathery top, grows in ground, is long, quite thin, crunchy, sweet
orange, grows in ground, is crunchy, good for eyes, enjoyed by rabbits, is grateable
orange, vegetable, hard, pointy, rabbits eat it, horses eat it, shave it
orange, long, thin, eatable, is green top
yellow, orange, feathery leaves, long, food, eyes
orange, crunchy, grows in soil, is edible
sprouts green hair, is orange, has ridges, grows in the ground, supposedly helps eyesight, can be used as dagger
is orange, is a vegetable, is long, looks like a penis, helps you see in the dark
is orange, is a root, is a vegetable, helps you see in the dark, grows
orange, cone shaped, long skinny, healthy, edible, comes from the ground

**BULL**

large, aggressive, masculine
angry, big, has horns, is stubborn, will chase you
big, hairy, horned, aggressive, proud
big, dangerous, has hooves, snorts, has a ring in nose, Taurus, brown
animal, is scary, hates red, is big, has horns, has a tail, has fur
has 4 legs, fights in Spain, is aggressive, has two horns, charges at the colour red
is large, is mean, associated with masculinity
angry, virile, ridden in rodeo, like a cow but with horns, dangerous in cartoons
has horns, gets angry, associated with red, strong
animal, fights in Spain, lured by red cloth, has horns, gets angry
is fierce, is an animal, likes red, has horns, is dangerous, base word for describing people who are rude
has horns, charges, male, large, four legs, thick hide, has a tail, hooves

**COW**

strong, large, heavy, grazes, loud, dairy resource, victim of humans for skin and meat
moo, black and white, big, fat, slow, milk, grazing
can be different colours, big, 4 legs, tail, produces milk, produces beef, eats grass

says moo, eats grass, produces milk, has 4 legs, bigger than me, produces pats
moos, farts a lot, produces milk, meat, makes patties, farm animal, 4 legs
is big, has 4 legs, is brown/white, eats grass, is soft, makes a moo sound, has a long tail
moos, gives milk, has calves, black and white, horns, lives in fields, eats
produces milk, is a large animal, has udders, farmed
beautiful, mother India, makes milk, tastes best medium rare, leather
black and white, moos, is cute, is stupid, makes milk, eats grass, big head
is big, produces milk, goes moo, gives birth, has four legs, coloured black and white
moos, chews cud, has four legs, makes milk, grazes, lives on a farm

**MASTER**
dominant, older, expert, masculine
in control, gives orders, has power
arrogant, pushy
has power, demands servitude, male, thick-set, invokes fear
is a teacher at school, type of degree, opposite to a servant
is powerful, is knowledgeable
comprehend, learn, understand a difficult subject, controls slaves
old word for teachers, in charge, owns slaves, implies subordination of others, counterpart of mistress
condescending, on top
in charge, sexual connotations, implies control
in charge, authority, high in hierarchical system
has control, is the original, dominates, is the largest, full cognition of

**APPRENTICE**
learner of a trade, undertakes training, vocational and professional qualifications
assistant, helper, training, learning
learning a trade, usually attached to master
TV show, lackey, makes coffee, learns his trade, will one day become master
squire, helper, novice, aide, does bitch work, in training
is a trainee, is learning
worker, could be hit in roman law to teach, skill
process to learn a trade, teaches a skill, usually a young person
takes about 3 years, lives on next to no money, starts life as a skivvy, should be good earner later, probably has gold earring
is a tv show, were abolished in the 80s and have been brought back, blacksmiths have them, work experience
helps someone, usually young, is in training
is young, learns, has a master

**TEA**
bitter, delicate, expensive, classy, asian, indian
hot, comforting, English, Indian, Chinese, in India very sweet, best from a pot
refreshment, made from leaves, made with water, hot, clear, comforting
is English, warm, comforting, brown, comes from India, lead to revolution in America, can be green & from China
opposite of coffee, is gross, is hot, is English
hot, comforting, soothes, calms my nerves, helps sleep, used in social customs, can be served cold, made from plant leaves & many varieties
breaks, relaxation, caffeine, British, tastes good

coffee's weaker younger brother, good, light brown, milky or not, made from leaves, from China or India, liked by the English, comes in clinky cups, gets bitter, sometimes hot, comforting

black, tasty, caffinated, herbal

is a drink, imported to England, hot or cold, usually hot, served with cookies sometimes

water infused with flavour, herbal or from tea tree, leaves may be roasted or green, black, green white, or from flowers, spices, bark, fruit, leaves

brewed herb, tranquilliser, drink always at hand, break in the day, gathering for refreshment

**COFFEE**

dark, soothing, aromatic, comforting drink for social occasions

black, energizing, good, aromatic

made from beans, very dark brown, usually hot, can be prepared in different ways, often drunk in the morning or after dinner

rich, smells nice, stains teeth, keeps you awake, gives you bad breath, makes you a snob

dark, caffeine, bitter, liquid, bean, morning, aroma

is brown, is fragrant, comes from coffee beans, is liquid

smell, served in cafes, made with coffee grounds

stimulant, something you drink, breakfast

wakes you up, makes lots of money for Starbucks, less money for the farmer, contains drugs, is brown, comes second to tea, helps sell houses

gives you bad breath, contains caffeine, sugar, coffee shops

is brown, smells strong, tastes bitter, made from beans

hot, can have sugar, can have milk, can be drunk, caffeinated

**QUEEN**

old, formal, rich, traditional

wears a crown, regal, wears hats, is posh

stately, refined, reserved, elegant, wise, hard-working

British, wears hats, goes racing, polite, gives weird gifts, likes colonies, is a character in Alice in Wonderland, many of them in other places

is a woman, opposite of a king, is a monarch, is rich, has jewellery, lives in a palace

is a diva, is dominant, is elegant, controls the land, wears tacky clothes, married to the king, rules over country

figurehead leader of England etc, royalty, has children, must pass on the line

England has one, wears a crown, outdated, rules, almost as good as a king, married to the king, sceptre, wears silly dresses, is on playing cards, whimsical, is gay

woman, wears crown, is a leader

woman ruler of a country, regal, wealthy, elegant, proper manners

female monarch, flamboyant cross-dresser, womanly power, remarkable clothing

bitchy man dressed as woman, female head monarch, high ranking card, female ruler, highly positioned bee

female ruler, E II current queen, married to king, usually very wealthy

**KING**

monarchy, ruler, ultimate power, overweight in stature, excessive wealth and status, not necessarily working for good of entire nation

wears a crown, powerful, man

head of monarchy, make, generally adult, has two legs, talks

wears a crown, uses a throne, gives orders, lives in a palace, is very sexually active

ruler, purple, royal, noble, giant ape, crown, has a queen

is royalty, has a crown, comes from monarchy

married to queen, wears a crown, head of state, earns respect
ruler, male monarch, ordained by god, head of state
has a kingdom, sometimes has a queen, sometimes is a queen, rules without merit, goes on stamps and cash, lives a life of luxury, runs the risk of deposement
rules countries, sells burgers, has lots of money, has a crown
runs a country, has a crown, commands an army, sits on a throne
wears a crown, rules the land, wields power, sits on a throne, lives in a castle

**CAT**
furry, smelly
soft, lazy, selfish, meows, eats birds, climbs trees, has kittens
pointy-eared, fluffy, chases mice, long-tailed, miows
warm, furry, purring, companion, independent, calm, home
an animal, has fur, meows, has a tail, good at jumping off high places
is independent, has a long tail, causes allergies in some people, chases mice, laps milk, has scratchy tongue, is a pet
not a dog, has 4 legs, meows, more of a loner pet, goes well with books and libraries, associated with single women
meows, is furry, smelly, soft, purrs, has claws, dog's enemy
soft, purrs, rubs on your leg, sleeps, climbs
animal, often a pet, likes to chase mice or string, nocturnal
four legs, glowing eyes, usually furry, lands on its feet, purrs
domestic feline, sun-seeking, client mistaken as pet, mouse catcher, tranquilliser, purr machine, refrigerator guardian

**DOG**
dirty animal, smelly, can be jovial or aggressive, saliva and noisy breathing, beautiful fur, alert, sometimes intelligent
barks, friendly, bites , eats, sleeps, loyal
4 legs, relatively small, generally furry, has a tail, barks, likes to run after balls, sticks, rabbits
has 4 legs, goes woof, chases their tail, chases cars, is stupid, loyal, bites people
barks, is loyal, good companion, fetches, has fur, four legs, puppy eyes
has 4 legs, barks, has tail, can be big or small, can be brown/black/white
not a cat, furry, bouncy, smelly when wet, friendly, pull on leash
man's best friend, barks, domesticated
loves unconditionally, barks, wags, eats, can be house-trained, may moult, may bite
loyal, smelly, hairy, obedient, makes a good friend
has four legs, barks/woofs, is loyal, has pups
has four legs, barks, is cute, pants, has a tail, has paws

**EMPLOYEE**
obedient, active
works, follows orders, earns money
busy, hassled, reluctant, looking forward to holidays
diligent, wears shirts, on time, drinks coffee, gets paid
opposite of boss, earns money, obeys orders
is hired by a boss, earns a pay cheque, is employed by a company, works for a living
part of a corporation, part of capitalism, gets paid regularly, works 9-5
works, wears uniform, refers to company as 'we'
works for someone else, wears uniform
subordinate, secure, works in cubicle, receives paychecks

worker within business, paid member within a labour division, pee-on, participating member of capitalist society, identified tax payer
earns money, is employed, has a boss

**EMPLOYER**
has power over you, dictates your duties, person with ultimate responsibility, should have expertise
boss, gives orders, manages people, observes, monitors
employs people to work, pays their wages, can refer to individual or corporation
pays money, gives orders, lays off workers, sends circular e-mails
the boss, owner, strict, stern, hires your, fires you, might help with problems
is older, earns more money, has responsibility
wages, good to work for
renumeration, can be single person or a company, responsibility
pays you, holds power over you, may not be as competent as you might assume, fills in your tax forms, might be contributing to the greater good, might be interested in employee's well being, is a person, too
wonderful, gives you money, generally a beurocratic moron
pays you, makes you work, runs a business
pays a salary, employs people

**HAND**
has 5 fingers, has nails, can move
has fingers & thumb, can be read, very useful, used for touching and writing, can be soft or rough
five fingered, dextrous, manipulative, touching, typing
has fingers, one of two, one is stronger, has veins, shows age, shows lifeline, used for tools
we have two, 5 fingers, nails, skin, necessary to function
claps, has a palm, has four fingers & thumb, has lines on palm, is important for accomplishing tasks
end of your arm, give over, 5 fingers
has fingers, lots of joints, types, plays instrument, covered in skin, veins, used to touch & hold, has nails, dexteruous
five fingers, palm
used for everything, in expressions such as lend a hand, has five fingers
five fingers, opposable thumbs, nails, knuckles, skin, creases on palms, visible veins, hair on knuckles
tool, help, fingers and opposing thumb, anatomical part, sensor, guide, can be opened wide, can be tightly shut

**GLOVE**
comforting, keeps hands warm, prevents frostbite or injury, can be glamorous
fits on hand, leather, fabric, protective, insulating
piece of clothing, protects hand, different materials, 4 fingers and thumb, can be protective for many reasons
has space for fingers, can be different materials, protects, keeps warm
fits your hand, five fingered, baseball, keeps you warm, leather, wool
has 5 areas for fingers, is shaped like a hand, made from material like leather
worn on hand, used to be made of kid leather, keep you warm, fluffy
clothing, keep you warm, can be knitted, come in a pair
fits, may have 5 fingers, may have no fingers, can be for warmth, anti-friction, anti-heat, loses feeling
is to hand as shoe is to foot, great in winter

goes on your hand, is warm, made of material, shaped like your hand, comes in different sizes
covers a hand, is warm

**MOTHER**
caring, nurturing, has children, cooks, listens
loving, caring, hard-working, busy
warm, gives comfort, gives nutrition, works, thinks, teaches, friend, is usually a bit weird
opposite of father, woman, older, caring, sharing, loving
comforts her children, is loving, is a woman, is strong, cares for her young, feeds the family
has children, loves, domesticity,
has a comforting bosom, has birthed
caring, gentle, affectionate
warm, caring, nurturing, centre of home, raises children
female, bears and/or raises children
vehicle into the world, nourisher, parent, smothers, source, antagonist, teacher
has a child, coming to visit soon, hard-working

**FATHER**
a catholic priest
man, old
has children, is male, human, has two legs, talks
provides, has fun with, teaches sport, is more distant
Dad, has kids, not the mother, stern, makes rules, strong, superman
is a man, is older, is caring
has children
patriarch, male parent, authoritarian
makes babies, might provide financial and emotional support, also used as a man for god or 'god's representatives', otusan in Japanese
good to have, mine is great
has children, is a man
is human, has two legs, had sex with mother to make child

**CUP**
has a handle, is round, holds water
used for holding liquids, good for eating cereal and ice cream, cylindrical in shape
vessel, holds tea & coffee, round, solid, hand-sized
crucible, to drink from, life, beaker, is round, is made of wood, is personal
holds fluid, plastic, glass, for drinks
container for liquid, unit of measurement, made of a variety of materials, glass, plastic, ceramic, is half full
saucer, drink out of, have tea in, hold hands under running water
holds liquid, round, smooth, shiney, not a mug because it doesn't have a handle
empty, full, vessel, container
vessel to drink from, cylindrical, measure in cooking
concave vessel, nor porous, holds liquid, usually stands on its own
container, holds, can be lifted, has a rim

**MUG**
porcelain, ceramic, sturdy large cup with handle, capable of holding very hot liquid
contains coffee, warm

implement used to drink out of, cylindrical usually, open top, handle, made out of ceramic, plastic or tin
has a handle, allows drinking of hot or cold drinks, displays humorous messages
cup, drink coffee in it, ceramic, has a handle, thick, breakable, picture from prison
is porcelain, has a handle, is round, has round edges
holds tea, warm, steam
breakable, tea, idiot, gullible
can be a face, or a drinks receptacle, best with tea inside, not too small, English homes have a random selection of them
funny joke to make, great with hot chocolate in, comforting, hate it when people use mine
made of ceramic, holds liquids, has a handle, withstands heat
holds liquid, is usually ceramic, has a handle

**DWARF**
short
short, big head
short, prejudice-fighting, bullied
is stunted, is a miner, wears hats, Tolkien, was used to teach me maths
small, Snow White, beard
short of stature, smaller than normal, is permanently small
short, not tall, tower over
small, wears hat, has beard, also real condition, jolly, saved snow white, whistles
short, characters in snow white
small person in clinical meaning, mythical creature, resides in wood, wears funny hats
short, some exaggerated features
smaller than average, features exaggerated, applied to plants & animals, friends of snow white

**GIANT**
huge in height or status, extremely tall or heavy person, or a company
green, scary, big, clumsy
man, two legs, exceptionally tall, mythical creature, extremely enormous
is big, attacks villages, carries club, lives in clouds accessible by beanstalk, treads on things
tall, big, larger than life
is big, is large, has two legs, has two arms
featured in Jack and the beanstalk, overlarge, mountains, very tall
large, ficticious, feature of children's fables, myth, male
brand for bicycle, or sweetcorn, enormous, lives at the top of beanstalk, adj or noun
Big, BFG, fairytales, children, dreams
much bigger than normal, taller, wider
is large, is human

**WIND**
cold, powerful
cold or hot, common in Wellington, makes umbrellas break, makes skirts fly up, good for sailing
breath, cool, exhilarating, oxygen
blows, is force, is strong, can calm, brings change, helps travel, brings omens
weather, cold, inconvenient
blows, rustles, lifts the sails, fills the lungs, whips, has direction, gusts,
willows, breeze, ambiguous concept

blows, is air, can be nice, rustles things, sometimes cold, caused by pressure changes, moves all over the world
breezy, light
weather phenomenon, rushing of air, soft or strong, warm or cold
moves, cools, picks up objects, carries them
a light zephyr, hurricane, air currents, invisible force

**RAIN**
replenishing, restorative, wet, makes land luscious
wet, tropical, cold, warm, refreshing
comes from condensation, made from H2O, falls to earth, is wet, can be cold, varying strengths, droplets
falls from sky, ruins drying, makes puddles, makes films atmospheric
it falls, is water, precipitation, wet, pours, thunder, storm
is wet, cold, comes in drops, makes puddles, comes from clouds
accompanied by wind, can be warm, wet, fine drizzle, soaking
wet, facet of weather, requires an umbrella, welly boots
falls, is wet, can be warm or cold, welcome or unwelcome, precipitation, falls at different angles, is a giver of life
never stops, is cold and damp, necessary to make stuff grow
comes from the sky, falls with gravity, is wet, made of water, usually cold, helps plants grow, comes from clouds
wet, falls from the sky, waters plants, comes from clouds

**MAN**
not a woman, no breasts, has a penis, has facial hair, has testosterone
gendered, diverse, human, idealist
wears a hat, wears a coat, separate to nature, is strong, leads
boy, human, animal
runs, has two legs, has two arms, has two feet, has a penis, produces testosterone, has skin
humanity, not woman, anthropology
male, has penis, is usually taller than women, what a piece of work is, can mean mankind
strong, stoic, firm
male person
mature male human, has penis and testicles, chest hair growth, facial hair, adam's apple, testosterone, deep voice
separate and part of nature, mortal, mammal on 2 legs
is male, is a human, is the opposite of woman

**WOMAN**
female in gender, mother, nurturer, strong, sensitive, independent, loving
mother, feminine, nurturing
female, human, two legs, talks, carries the young
can multitask, faces discrimination, has children, harder working, less fun
sexy, breast, is not a man, smells nice, soft skin, what a man needs, makes babies
has two breasts, a vagina, can have children
often carrying bags, high heels, some wear smart suits, others don't
femininity, motherhood, childrearing, beauty, other
life-giver, pulse-quickener, beautiful, fertile, amazing skin, smells good, can taste amazing
beautiful
adult, can give birth, has two legs
is human, has breasts, can give birth

**HILL**
raised earth, nice for rolling down, not nice for walking up, good views from the top
green, grass, high up
steep, smaller than mountain, grassy, occurs in nature
is steeper than flat ground, can be covered in grass, can obscure vision, can be used as look-out, can be steep
not as tall as a mountain, sloping, hard to climb up
hard to climb, roundish, green, not quite mountain
grass, ants
smaller than mountain, inclined part of terrain
raised, gentle, curved, water runs off it
slope, incline up and downs, something to climb up and roll down, red earth and stone, green grass and flower, view point
is something to run up, smaller than a mountain
small mountain

**VALLEY**
between hills or in forests, beautiful, natural, green, fresh, soggy underfoot
green, rolling hills
dip between hills, often caused by river
in between mountains, contains Welsh people, contains LA airheads, has water at the bottom
shallow, between mountains, low lying
is low, is near something higher, is grassy, is a dip
with rivers, near hills, green plants, cows and sheep, can be attractive
hillside, mountains, trees, countryside
lies between, was carved by water, sometimes green, divides hills and mountains
is green land between two hills, deep, often people live in it
between two hills, goes down, usually has a river at the bottom
is a dent in the landscape, is below surrounding land

**TEACHER**
dominant, leads class, assigns work, sets agenda
one who instructs, a knowledgeable person
imparts knowledge, human, wise, has students, should be patient
helps others learn, is kind (ideally), is patient, knows more than the student, grades papers, works hard
passes on knowledge, has authority, has mastered a subject
teaches, wears glasses, conveys knowledge, guides study, has chalk board, is poorly paid
in charge
someone who instructs
human, educational authority
guide, coach, reveals knowledge or skill, mentor, person or experience
helps people learn, teaches in classroom
instructs students, leads education

**STUDENT**
learner, hard-working, curious, driven to achieve good results
poor
human, male or female, talks, two legs, studying at school, often inebriated
doesn't work, sleeps in, wastes time, protests, watches TV, makes mess, spends money
pupil, learner, is in school, is taught, studies, works hard,

is a person who goes to school, studies, writes essays, takes exams, goes to class, spends time at the library, working towards a degree

goes to university or school, reads books, usually young person

learning, reading books, writing essays, taking exams

thinks, learns, drinks, worries, seeks to embetter themselves and the world they live in, to be filled by self and others with information and interpretations, seen by some as bludgers

poor, free, drunk, young, fun

is lazy or studious, learns, parties, young

is human, goes to school, takes classes, learns, has teachers

**Appendix 9** T-score and GOE-rating for all German pairs

| Word 1 | Freq | Word 2 | Freq | Total FoC | T-Score | QR 1-2 | QR 2-1 | QR |
|---|---|---|---|---|---|---|---|---|
| unten | 95453 | oben | 207252 | 15139 | 122.689 | 1.03 | 1.03 | 1.03 |
| hell | 19809 | dunkel | 32350 | 1138 | 33.692 | 1.05 | 1.03 | 1.04 |
| alt | 338295 | neu | 505688 | 3473 | 52.598 | 1.08 | 1.03 | 1.055 |
| groß | 346812 | klein | 195763 | 15599 | 123.709 | 1.08 | 1.08 | 1.08 |
| heiß | 53896 | kalt | 48404 | 954 | 30.702 | 1.08 | 1.08 | 1.08 |
| jung | 155435 | alt | 338295 | 39087 | 197.123 | 1.13 | 1.03 | 1.08 |
| lang | 423083 | kurz | 578171 | 6815 | 76.087 | 1.13 | 1.08 | 1.105 |
| falsch | 92620 | richtig | 342200 | 2852 | 52.109 | 1.08 | 1.15 | 1.115 |
| langsam | 120305 | schnell | 446308 | 1146 | 30.391 | 1.2 | 1.03 | 1.115 |
| schwer | 412910 | leicht | 314557 | 5251 | 68.552 | 1.13 | 1.13 | 1.13 |
| hart | 101564 | weich | 12530 | 596 | 24.299 | 1.13 | 1.19 | 1.16 |
| stark | 400220 | schwach | 41859 | 1127 | 32.481 | 1.05 | 1.3 | 1.175 |
| ungerade | 831 | gerade | 646331 | 173 | 13.063 | 1.15 | 1.23 | 1.19 |
| leicht | 314557 | schwierig | 90179 | 186 | 9.099 | 1.25 | 1.15 | 1.2 |
| schwierig | 90179 | einfach | 586074 | 559 | 18.765 | 1.23 | 1.18 | 1.205 |
| illegal | 27150 | legal | 10207 | 396 | 19.869 | 1.15 | 1.27 | 1.21 |
| gut | 1430211 | schlecht | 157271 | 6071 | 71.617 | 1.13 | 1.3 | 1.215 |
| gesund | 46397 | krank | 41996 | 653 | 25.387 | 1.23 | 1.21 | 1.22 |
| sauber | 34108 | schmutzig | 3503 | 38 | 6.122 | 1.23 | 1.23 | 1.23 |
| weiblich | 13540 | männlich | 11944 | 2663 | 51.597 | 1.28 | 1.18 | 1.23 |
| böse | 48416 | gut | 1430211 | 5008 | 68.632 | 1.38 | 1.15 | 1.265 |
| erstes | 97737 | letztes | 60296 | 215 | 13.785 | 1.15 | 1.38 | 1.265 |
| glücklich | 76882 | unglücklich | 24915 | 191 | 13.517 | 1.3 | 1.25 | 1.275 |
| Zwerg | 4903 | Riese | 7462 | 140 | 11.825 | 1.18 | 1.39 | 1.285 |
| ordentlich | 39576 | unordentlich | 397 | 7 | 2.632 | 1.25 | 1.33 | 1.29 |
| unglücklich | 24915 | glücklich | 76882 | 191 | 13.517 | 1.25 | 1.33 | 1.29 |
| feminin | 912 | maskulin | 316 | 39 | 6.244 | 1.35 | 1.24 | 1.295 |
| Tod | 225707 | Leben | 798017 | 12762 | 109.49 | 1.33 | 1.3 | 1.315 |
| irrational | 888 | rational | 2691 | 24 | 4.897 | 1.2 | 1.45 | 1.325 |
| einfach | 586074 | schwer | 412910 | 1200 | 19.398 | 6.65 | 6.6 | 1.34 |
| riesig | 17167 | winzig | 2876 | 29 | 5.365 | 1.25 | 1.43 | 1.34 |
| anwesend | 31933 | abwesend | 3661 | 37 | 6.04 | 1.28 | 1.52 | 1.4 |
| geben | 729005 | nehmen | 478566 | 3433 | 45.599 | 1.58 | 1.23 | 1.405 |
| moralisch | 8193 | unmoralisch | 1686 | 22 | 4.683 | 1.33 | 1.48 | 1.405 |
| hinein | 56573 | hinaus | 256608 | 406 | 18.577 | 1.4 | 1.43 | 1.415 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| unmöglich | 49085 | möglich | 581288 | 256 | 12.108 | 1.38 | 1.45 | 1.415 |
| schlafend | 1548 | wach | 20696 | 10 | 3.14 | 1.35 | 1.5 | 1.425 |
| passend | 24561 | unpassend | 2002 | 12 | 3.433 | 1.35 | 1.55 | 1.45 |
| unloyal | 22 | loyal | 2683 | 0 | 0 | 1.5 | 1.4 | 1.45 |
| spät | 104132 | früh | 178181 | 2454 | 48.72 | 1.48 | 1.45 | 1.465 |
| wahr | 54270 | falsch | 92620 | 379 | 18.904 | 1.53 | 1.42 | 1.475 |
| betrunken | 14281 | nüchtern | 13518 | 43 | 6.493 | 1.38 | 1.58 | 1.48 |
| hungrig | 5771 | satt | 15322 | 44 | 6.604 | 1.23 | 1.73 | 1.48 |
| direkt | 249382 | indirekt | 21595 | 3250 | 56.802 | 1.55 | 1.45 | 1.5 |
| Mann | 835267 | Frau | 844270 | 37543 | 185.818 | 2.95 | 2.63 | 1.5 |
| anziehen | 15962 | ausziehen | 6237 | 38 | 6.129 | 1.48 | 1.55 | 1.515 |
| desinteressiert | 874 | interessiert | 100522 | 3 | 1.621 | 1.4 | 1.64 | 1.52 |
| vor | 6365082 | hinter | 517920 | 14426 | 60.22 | 1.48 | 1.58 | 1.53 |
| ungesund | 2186 | gesund | 46397 | 64 | 7.972 | 1.5 | 1.58 | 1.54 |
| verbieten | 18367 | erlauben | 29237 | 70 | 8.226 | 1.6 | 1.48 | 1.54 |
| uninteressiert | 364 | interessiert | 100522 | 1 | 0.92 | 1.68 | 1.45 | 1.565 |
| Berg | 115945 | Tal | 54716 | 1049 | 31.96 | 1.4 | 1.79 | 1.595 |
| verschlossen | 12018 | offen | 257267 | 69 | 7.494 | 1.38 | 1.82 | 1.6 |
| antworten | 57656 | fragen | 405353 | 13965 | 117.742 | 1.65 | 1.63 | 1.64 |
| runter | 18644 | hoch | 325231 | 904 | 29.626 | 1.7 | 1.65 | 1.675 |
| real | 59008 | irreal | 614 | 27 | 5.18 | 1.48 | 1.88 | 1.68 |
| dienstlich | 2300 | privat | 63078 | 426 | 20.624 | 1.9 | 1.58 | 1.74 |
| verkaufen | 97557 | kaufen | 105307 | 1994 | 44.152 | 3.75 | 4.55 | 1.74 |
| unter | 3576918 | über | 5780459 | 52028 | 30.308 | 1.93 | 1.55 | 1.75 |
| nah | 41479 | weit | 438067 | 280 | 14.363 | 1.68 | 2.06 | 1.87 |
| akkurat | 2050 | inakkurat | 3 | 0 | 0 | 1.9 | 1.85 | 1.875 |
| Feuer | 144038 | Wasser | 351446 | 3327 | 55.765 | 2 | 1.75 | 1.875 |
| kommen | 917319 | gehen | 664390 | 6581 | 64.73 | 1.53 | 2.24 | 1.885 |
| alleine | 96405 | zusammen | 639483 | 401 | 13.307 | 1.93 | 1.95 | 1.94 |
| miserabel | 3444 | exzellent | 4831 | 0 | 0 | 1.83 | 2.05 | 1.94 |
| unverheiratet | 941 | verheiratet | 44182 | 20 | 4.451 | 1.93 | 1.97 | 1.95 |
| angreifen | 10928 | verteidigen | 42699 | 81 | 8.886 | 2.13 | 1.83 | 1.98 |
| unverschlossen | 578 | verschlossen | 12018 | 0 | 0 | 1.75 | 2.21 | 1.98 |
| weich | 12530 | fest | 664572 | 58 | 5.23 | 1.85 | 2.13 | 1.99 |
| süß | 12943 | sauer | 45217 | 350 | 18.64 | 2.13 | 1.88 | 2.005 |
| einfarbig | 566 | bunt | 36213 | 44 | 6.626 | 2.05 | 1.98 | 2.015 |
| mehr | 4011482 | weniger | 602034 | 68504 | 241.599 | 1.85 | 2.24 | 2.045 |
| Soll | 27880 | Haben | 46231 | 433 | 20.673 | 2.48 | 1.63 | 2.055 |
| ungehorsam | 1678 | brav | 15829 | 3 | 1.698 | 1.78 | 2.38 | 2.08 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| miteinbeziehen | 1088 | ausschließen | 17967 | 1 | 0.957 | 1.8 | 2.39 | 2.095 |
| fest | 443728 | weich | 12530 | 58 | 6.022 | 2.2 | 2.15 | 2.175 |
| unreal | 121 | real | 59008 | 3 | 1.723 | 2.5 | 1.94 | 2.22 |
| einflussreich | 411 | einflusslos | 41 | 0 | 0 | 2.33 | 2.12 | 2.225 |
| privat | 63078 | staatlich | 16988 | 119 | 10.694 | 2.3 | 2.21 | 2.255 |
| duzen | 900 | siezen | 245 | 50 | 7.07 | 2.4 | 2.28 | 2.34 |
| exakt | 42614 | unpräzise | 892 | 0 | 0 | 2.43 | 2.39 | 2.41 |
| amoralisch | 99 | moralisch | 8193 | 2 | 1.412 | 2.38 | 2.52 | 2.45 |
| vermieten | 6773 | mieten | 21597 | 39 | 6.193 | 2.6 | 2.3 | 2.45 |
| tugendhaft | 171 | sündig | 207 | 0 | 0 | 2.45 | 2.5 | 2.475 |
| riesig | 17167 | klein | 195763 | 48 | 5.869 | 2.28 | 2.73 | 2.505 |
| Vermieter | 20356 | Mieter | 48521 | 2380 | 48.741 | 2.68 | 2.35 | 2.515 |
| Land | 731001 | Stadt | 1336827 | 25911 | 147.722 | 2.45 | 2.61 | 2.53 |
| Henne | 3932 | Hahn | 86742 | 184 | 13.509 | 2.55 | 2.58 | 2.565 |
| schlecht | 157271 | exzellent | 4831 | 5 | 1.494 | 2.45 | 2.8 | 2.625 |
| verschieden | 13648 | ähnlich | 142149 | 67 | 7.668 | 2.85 | 2.53 | 2.69 |
| richtig | 342200 | inkorrekt | 145 | 4 | 1.945 | 2.38 | 3.03 | 2.705 |
| mild | 6008 | würzig | 1024 | 32 | 5.654 | 4.38 | 4.75 | 2.725 |
| winzig | 2876 | groß | 346812 | 25 | 4.564 | 2.48 | 2.97 | 2.725 |
| gigantisch | 3267 | mini | 9531 | 0 | 0 | 2.2 | 3.27 | 2.735 |
| kühl | 22800 | warm | 34646 | 178 | 13.212 | 2.48 | 3.03 | 2.755 |
| reparieren | 9775 | beschädigen | 3731 | 2 | 1.357 | 2.68 | 2.85 | 2.765 |
| lernen | 211052 | lehren | 16997 | 825 | 28.45 | 3.13 | 2.45 | 2.79 |
| Mann | 835267 | Weib | 4340 | 245 | 15.147 | 1.48 | 1.52 | 2.79 |
| Vergangenheit | 163546 | Gegenwart | 46635 | 5895 | 76.562 | 3.35 | 2.3 | 2.825 |
| Opfer | 254529 | Mörder | 28139 | 638 | 24.639 | 2.93 | 2.73 | 2.83 |
| kochend | 307 | eisig | 1830 | 0 | 0 | 2.63 | 3.08 | 2.855 |
| Mama | 32248 | Papa | 28595 | 4927 | 70.163 | 2.78 | 2.97 | 2.875 |
| Lehrer | 199088 | Schüler | 566654 | 28444 | 167.193 | 3.23 | 2.58 | 2.905 |
| Schwester | 84965 | Bruder | 150703 | 3054 | 54.757 | 2.8 | 3.03 | 2.915 |
| Zukunft | 518747 | Gegenwart | 46635 | 4044 | 62.762 | 2.98 | 2.85 | 2.915 |
| Doktor | 17981 | Patient | 31282 | 49 | 6.824 | 2.98 | 2.94 | 2.96 |
| Meister | 203124 | Lehrling | 12601 | 269 | 16.06 | 3.18 | 2.91 | 3.045 |
| Angestellter | 12746 | Chef | 198590 | 78 | 8.206 | 3.03 | 3.12 | 3.075 |
| Vergnügen | 38042 | Arbeit | 775269 | 801 | 26.028 | 3.35 | 2.88 | 3.115 |
| Mutter | 397872 | Vater | 310066 | 31909 | 177.123 | 3.43 | 2.93 | 3.18 |
| daheim | 51420 | weg | 180563 | 184 | 12.071 | 3.08 | 3.33 | 3.205 |
| verheiratet | 44182 | single | 13323 | 26 | 4.847 | 2.28 | 2.53 | 3.277 |
| Bulle | 4510 | Kuh | 18761 | 34 | 5.799 | 2.98 | 3.58 | 3.28 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| bitter | 27780 | süß | 12943 | 251 | 15.793 | 3.18 | 3.39 | 3.285 |
| warm | 34646 | frostig | 2029 | 9 | 2.948 | 3.23 | 3.48 | 3.355 |
| kühl | 22800 | heiß | 53896 | 98 | 9.628 | 3.18 | 3.55 | 3.365 |
| ganz | 1378076 | halb | 49231 | 421 | 13.303 | 3.7 | 3.09 | 3.395 |
| Festnetz | 10008 | Handy | 54734 | 379 | 19.406 | 3.6 | 3.21 | 3.405 |
| früher | 290291 | gegenwärtig | 37991 | 39 | 2.391 | 3.33 | 3.48 | 3.405 |
| geschieden | 8341 | verheiratet | 44182 | 250 | 15.76 | 3.15 | 3.7 | 3.425 |
| jagen | 13999 | fliehen | 12271 | 4 | 1.812 | 3.33 | 3.55 | 3.44 |
| auspacken | 2868 | packen | 22816 | 5 | 2.172 | 3.7 | 3.33 | 3.515 |
| falsch | 92620 | passend | 24561 | 4 | -0.481 | 3.2 | 3.85 | 3.525 |
| Student | 23338 | Dozent | 9346 | 33 | 5.661 | 3.98 | 3.08 | 3.53 |
| frostig | 2029 | sommerlich | 2322 | 0 | 0 | 3.43 | 3.64 | 3.535 |
| bunt | 36213 | unbunt | 26 | 1 | 0.997 | 3.58 | 3.68 | 3.63 |
| Katze | 31395 | Hund | 87212 | 4042 | 63.482 | 4.4 | 2.88 | 3.64 |
| woher | 28444 | wohin | 47645 | 913 | 30.118 | 3.58 | 3.7 | 3.64 |
| Eltern | 525084 | Kinder | 1499055 | 66010 | 250.239 | 3.5 | 3.79 | 3.645 |
| Königin | 41606 | König | 138165 | 2169 | 46.303 | 4.08 | 3.23 | 3.655 |
| Tante | 21838 | Onkel | 25000 | 1050 | 32.366 | 4.1 | 3.33 | 3.715 |
| einfach | 586074 | doppelt | 61861 | 192 | 8.147 | 1.28 | 1.4 | 3.72 |
| Nichte | 5991 | Neffe | 5961 | 119 | 10.901 | 3.65 | 3.91 | 3.78 |
| Maus | 26994 | Katze | 31395 | 1816 | 42.571 | 3.98 | 3.65 | 3.815 |
| dick | 34756 | schmal | 6537 | 7 | 2.458 | 3.43 | 4.21 | 3.82 |
| scharf | 46067 | fad | 2931 | 3 | 1.561 | 3.53 | 4.18 | 3.855 |
| lernen | 211052 | verlernen | 568 | 23 | 4.741 | 3.45 | 4.3 | 3.875 |
| modisch | 3177 | traditionell | 46543 | 6 | 2.317 | 3.48 | 4.33 | 3.905 |
| heute | 1998661 | morgen | 485703 | 45963 | 204.509 | 4.65 | 3.55 | 4.1 |
| ausleihen | 6221 | verleihen | 24602 | 4 | 1.833 | 4.3 | 4 | 4.15 |
| verkaufen | 97557 | behalten | 64577 | 117 | 9.545 | 1.93 | 1.55 | 4.15 |
| unscheinbar | 2567 | kolossal | 480 | 0 | 0 | 4.15 | 4.25 | 4.2 |
| farbig | 5412 | weiß | 468495 | 68 | 7.575 | 4.38 | 4.03 | 4.205 |
| Fleisch | 54646 | Fisch | 34271 | 3185 | 56.363 | 4.45 | 4.15 | 4.3 |
| Hexe | 10892 | Fee | 5490 | 43 | 6.537 | 4.15 | 4.55 | 4.35 |
| grün | 65586 | rot | 77043 | 5491 | 73.952 | 4.43 | 4.28 | 4.355 |
| zufriedenstellend | 10261 | schlecht | 157271 | 17 | 3.269 | 4.2 | 4.7 | 4.45 |
| Schauspielerin | 37777 | Schauspieler | 93643 | 425 | 20.241 | 4.45 | 4.48 | 4.465 |
| mild | 6008 | kalt | 48404 | 39 | 6.143 | 2.55 | 2.9 | 4.565 |
| dunkel | 32350 | blass | 5677 | 4 | 1.799 | 4.4 | 4.76 | 4.58 |
| durstig | 1870 | hungrig | 5771 | 588 | 24.247 | 4.85 | 4.42 | 4.635 |
| stehlen | 12546 | spenden | 133234 | 3 | -0.373 | 3.98 | 5.35 | 4.665 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| gegenüber | 478057 | neben | 842634 | 921 | 1.385 | 4.33 | 5.03 | 4.68 |
| Tee | 25809 | Kaffee | 117728 | 4826 | 69.373 | 4.88 | 4.48 | 4.68 |
| brütend | 351 | kühl | 22800 | 1 | 0.982 | 4.73 | 4.68 | 4.705 |
| mangelhaft | 8470 | angemessen | 22233 | 0 | 0 | 4.78 | 4.7 | 4.74 |
| freundlich | 33820 | kalt | 48404 | 45 | 6.175 | 4.75 | 4.8 | 4.775 |
| Mutter | 397872 | Tochter | 236855 | 12546 | 110.173 | 4.75 | 4.88 | 4.815 |
| brilliant | 313 | mittelmäßig | 1378 | 0 | 0 | 4.98 | 4.75 | 4.865 |
| akzeptabel | 10381 | furchtbar | 8550 | 0 | 0 | 4.58 | 5.21 | 4.895 |
| hilflos | 15371 | hilfreich | 20831 | 3 | 1.328 | 4.05 | 5.58 | 4.94 |
| hilfreich | 20831 | hilflos | 15371 | 3 | 1.328 | 5.3 | 4.58 | 4.94 |
| krabbeln | 3062 | gehen | 664390 | 35 | 5.165 | 5.03 | 5.28 | 5.155 |
| Besitzer | 85052 | Bewohner | 115822 | 86 | 6.955 | 5.08 | 5.4 | 5.24 |
| praktisch | 97602 | formal | 10358 | 18 | 3.722 | 5.05 | 5.43 | 5.24 |
| frisch | 79562 | tot | 60963 | 9 | -0.527 | 5.33 | 5.48 | 5.405 |
| Tochter | 236855 | Vater | 310066 | 8898 | 92.63 | 5.7 | 5.15 | 5.425 |
| orange | 16996 | blau | 36757 | 802 | 28.271 | 5.38 | 5.55 | 5.465 |
| süß | 12943 | trocken | 39862 | 54 | 7.195 | 5.58 | 5.43 | 5.505 |
| blau | 36757 | gelb | 26569 | 3384 | 58.135 | 5.6 | 5.42 | 5.51 |
| Huhn | 9460 | Ei | 25423 | 290 | 16.998 | 5.73 | 5.3 | 5.515 |
| mittelmäßig | 1378 | gut | 1430211 | 90 | 9.033 | 5.6 | 5.64 | 5.62 |
| jagen | 13999 | ausbrechen | 5272 | 1 | 0.838 | 5.88 | 5.38 | 5.63 |
| erwerben | 38899 | handeln | 93991 | 16 | 2.005 | 5.6 | 5.78 | 5.69 |
| schmal | 6537 | lang | 423083 | 196 | 13.568 | 6.35 | 5.52 | 5.935 |
| ausgeben | 21255 | kaufen | 105307 | 35 | 5.09 | 6.08 | 5.91 | 5.995 |
| unmoralisch | 1686 | amoralisch | 99 | 5 | 2.235 | 6.15 | 5.91 | 6.03 |
| rufen | 52037 | brüllen | 3188 | 14 | 3.644 | 5.93 | 6.2 | 6.065 |
| Seite | 732578 | Buch | 250102 | 792 | 13.937 | 6.03 | 6.12 | 6.075 |
| lau | 6383 | mild | 6008 | 3 | 1.683 | 6.28 | 6 | 6.14 |
| Entgegnung | 750 | Antwort | 144522 | 2 | 1.246 | 6.23 | 6.1 | 6.165 |
| Hand | 354194 | Finger | 46930 | 1445 | 37.059 | 6.13 | 6.24 | 6.185 |
| identisch | 10270 | ähnlich | 142149 | 49 | 6.544 | 6.28 | 6.09 | 6.185 |
| schlecht | 157271 | falsch | 92620 | 293 | 15.26 | 6.3 | 6.12 | 6.21 |
| befolgen | 3299 | einwilligen | 510 | 0 | 0 | 6.08 | 6.39 | 6.235 |
| Erbse | 1911 | Karotte | 997 | 1 | 0.995 | 6.43 | 6.09 | 6.26 |
| Fuß | 118459 | Bein | 31334 | 199 | 13.532 | 6.35 | 6.21 | 6.28 |
| Regen | 105673 | Wind | 107424 | 3759 | 60.906 | 6.2 | 6.38 | 6.29 |
| Tasse | 9480 | Becher | 16705 | 27 | 5.129 | 6.23 | 6.36 | 6.295 |
| lang | 423083 | viereckig | 201 | 2 | 1.283 | 6.53 | 6.09 | 6.31 |
| Henkel | 15598 | Tasse | 9480 | 19 | 4.284 | 6.45 | 6.23 | 6.34 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| fremd | 18165 | kalt | 48404 | 24 | 4.507 | 6.4 | 6.35 | 6.375 |
| wütend | 11316 | verrückt | 20363 | 4 | 1.748 | 6.35 | 6.45 | 6.4 |
| ordentlich | 39576 | sauber | 34108 | 364 | 18.924 | 6.55 | 6.33 | 6.44 |
| helle | 20478 | clever | 9949 | 0 | 0 | 6.55 | 6.58 | 6.452 |
| höflich | 8056 | kurz | 578171 | 38 | 4.515 | 6.2 | 6.82 | 6.51 |
| verrucht | 327 | scheußlich | 709 | 0 | 0 | 6.48 | 6.55 | 6.515 |
| unglücklich | 24915 | traurig | 32588 | 28 | 4.956 | 6.75 | 6.36 | 6.555 |
| weg | 180563 | tot | 60963 | 112 | 8.313 | 6.55 | 6.58 | 6.565 |
| einfach | 586074 | weich | 12530 | 75 | 6.81 | 3.83 | 3.61 | 6.625 |
| still | 37726 | ruhig | 64777 | 252 | 15.538 | 6.55 | 6.7 | 6.625 |
| grau | 19415 | alt | 338295 | 160 | 11.516 | 6.7 | 6.63 | 6.665 |
| Blume | 10657 | Rose | 36795 | 80 | 8.848 | 6.6 | 6.76 | 6.68 |
| König | 138165 | Krone | 64009 | 297 | 16.113 | 6.68 | 6.7 | 6.69 |
| Hund | 87212 | Tier | 79783 | 666 | 25.218 | 6.85 | 6.83 | 6.84 |
| klebrig | 544 | kalt | 48404 | 3 | 1.698 | 6.85 | 6.97 | 6.91 |

## **Appendix 10** German attribute listing results

**ERBSE**
Hülsenfrucht, Feldgewächs, grün, frisch oder getrocknet gut
klein, grün, gesund, tiefgekühlt
rund, grün, ekelig
klein, grün, essbar, Gemüse, rund, viel Arbeit beim "putzen", gesund, weich, lecker
grün, rund, weich
grün, rund, viereckig, klein, Biologie
klein, grün, kann man essen
grün, klein, rund, weich
grün, rund, tiefgekühlt, wächst in einer Schote, wird gepult
rund, grün, essen, Schote, Wurm
grün, klein, hart, dampfend, weichgekocht
grün, verursacht Blähungen, ist rund
rund, grün, ist eine Hülsenfrucht, gibt es in der Dose
grün, gesund, lecker, Kinder mögen sie selten, ist Gemüse, in einer Hülse
grün, klein, rund, lecker, ein Gemüse

**KAROTTE**
orange, lang, hat grünes Gebinde, heilt Augen
orange, gesund, lecker
orange, lecker, Hase, Gemüse, Wurzel, Salat, Augen
orange, wächst unter der Erde, ist gut für die Augen, ist mit Hasen assoziiert, Carotin, grüne Blätter, ist ein Gemüse
orange, wächst im Boden, mögen Hasen
orange, länglich, Erdgewächs
Essen, Gemüse, Salat, schmeckt gut, orange, lecker
gesund, gut für die Augen, ist orange, ist Rohkost, schmeckt roh und gekocht
orange, gesund, knackig, lecker, gut für die Augen
orange, lang, Gemüse, gesund, ist ein Kinderessen, hat einen süßen Kern
orange, auch Möhre, gelbe Rübe, Hase, Felt, Grünzeug, gesund
orange, schmeckt lecker, vor allem der Saft, wächst in der Erde im Boden, sollte man schälen
orange, lang, spitz, grün am Ende
orange, lang, gesund, Vitamin A, schmeckt nicht, Karottensaft

**BULLE**
männlich, bovine, hat Hörner, kann auch als Schimpfwort verwendet werden
Stier, redBull, riesig, stark
massig, braun, hat Hörner, gibt keine Milch, kräftig
Spanien, Tier, groß, dick, braun, gefährlich, Synonym für Polizist, dann eher Schimpfwort
Buckel, Nacken, Hörner
Polizist, Stier, rot, Spanien
groß, kräftig, immer nur einen pro Bauernhof, gefährlich, aggressiv, ugs für Polizist
groß, gefährlich, hat Hörner, isst Grass, hat Hufe, schnaubt, hat Fell, hat einen Schwanz
anderes Wort für Polizist, steht an der Frankfurter Börse, ist kraftvoll, hat Hörner, ist ein einziger kompakter Muskel
Tier, männlich, Hörner, Penis, Aggression, Kuh
hat Hörner, hat Hoden, ist stark, hat glänzendes Fell, ist schnell

ist ein kastrierter Stier, hat Hörner
Rind, hat Ring durch die Nase, schnaubt, groß, macht Angst
zweideutig, hat Hörner, ist stark
groß, stark, wild, hat Hörner, ist ein Säugetier, männlich

**KUH**
ist fleckig, hat großes glitschiges Maul, Kuhdung
muh, milchgebend, stinkt, dick
Milch, Alm, Weide, Berge, Milka, klatschen, Kuhfladen
gibt Milch, grast auf Weiden, ist ein Widerkäuer, weibliches Rindvieh, hat Euter, ist weiss mit schwarzen Flecken, Logo von la Vache Qui Rit
macht Muh, gibt Milch, hat Hörner, frisst Gras
gibt Milch, Bauernhof, braun, macht Muh
Rind, gibt Milch und Fleisch, ist ein Schimpfwort, gefleckt, Milka , lila
gibt Milch, ist braun, gescheckt, schöne Augen, ist ein Wiederkäuer, frisst Gras, muht, trägt oft eine Glocke um den Hals, lebt auf der Alm
heilig, gibt Milch, frisst Gras, macht muh, wiederkäuend, Milka
hat Flecken, kann man essen, ist auf der Wiese
gefleckt, Tier, Weide, Milch, Euter, Rindfleisch, Kalb, Dumm, lebt im Stall
gibt Milch, gibt Fleisch, gebärt Kälber, ist manchmal lila
braun, gibt Milch, macht Mist, meistens auf Weiden
gibt Milch, ist gefleckt, steht auf der Bergwiese, hat eine Kuhglocke

**MEISTER**
machts vor, kanns am besten, hat das sagen, betreut Lehrlinge
lernen, lehren, erfahren
ist sehr gut, wird zu Hilfe gerufen, hat gewonnen, besitzt eine Medallie, trainiert sich,
Gewinner, Sieger, Fußball, Jesus Christus, unser Herr und Meister, Erlöser
Judo, Mann, etwas können
Lehrer, Jesus, Chef
Lehrer, weiss immer was, intelektuell überlegen
ist erfahren, ist kein Geselle, übt ein Handwerk aus, kann was er tut gut
kommt aus Fernost, lehrt Handwerk, lehrt Zauberei, hat einen feinen grauen Bart, trinkt Bier, ist ruppig
Lampe, Autorität, Handwerk, Sport
hat Schüler, weiss Bescheid, lehrt andere
ist ein Spezialist auf seinem Gebiet
Werkstatt, ist Lehrer, hat Autorität
handwerklich begabt
ist weise, klug, erfahren, mit Respekt zu behandeln

**LEHRLING**
hat unangenehme Arbeiten zu verrichten, ist jung, ist männlich
handwerklich, körperlich anstrengend, jugendlich
Ausbildung, Hauptschule, Kaffee, Autos, Handwerk
ist in der Ausbildung, hat einen Meister, lernt einen handwerklichen Beruf
lernt einen Beruf, ist jung, verdient kaum Geld
Betrieb, Berufsschule, schreibt Schulaufgaben
Azubi, Lehrjahre sind keine Herrenjahre, lernen, Pauken, büffeln, Prüfungen, lange her
ist jung, geht zur Berufsschule, muss noch viel lernen, wird eine Prüfung schreiben, durchläuft viele Abteilungen
jung, Auszubildender, lernwillig

hat einen Meister, ist ein Anfänger, jung
Auszubildender, Stift, Zauberlehrling, lernen, Lehrer, Berufsschule, Zeugnis, Arbeit
Azubi, Lehrjahre, Lehrgeld, jung, frech
arm, jung, Hauptschulabschluss
jung, trägt Blaumann, repariert Autos

**TEE**
bringt mich durch den Tag, ist erholsam, fruchtig, heiss oder kalt trinkbar
Wasser mit wenig Geschmack, ungesüßt, unhandlicher Beutel
flüssig, heiß, gesund, wird in Krankheit oder als Brauch getrunken, vielfältige Geschmacksrichtungen
Getränk, gesund, Wasser, viele verschiedene Sorten, Indien, heiß und kalt genießbar, ziehen lassen
gut, Milch, heiß
diverse Geschmacksrichtungen, Gesund, Zucker, Glas , Löffel
trinkbar, gut, Grün, Schwarz Kräuter, Alternative zu Kaffee
schmekt nicht, ist eine Pflanze, ist grün, hält einen Golfball, ist wohl gesund
lässt man ziehen, wird mit heißem Wasser übergossen, gibt es auch granuliert & löslich, wird in England getrunken, ist gemütlich
Wasser, trinken, heiß
grüner, schwarzer, weisser Tee, ist aus Kräutern, ist gesund, ist warm
ist heiß, kalorienarm, koffeinhaltig
schwarz, Indien, heiß, mögen viele lieber als Kaffee
schwarz, grün, Pendant zu Kaffee, muss ziehen, ist heiss
ist heiß, gibt es in verschiedenen Sorten, wird viel in England getrunken, manchmal mit Milch, schwarzer und grüner Tee enthalten Koffein

**KAFFEE**
ist schwarz, heiß, am Besten schwarz
schwarz, aromatisch, koffeinhaltig
überleben, müde, braucht, Milch, Zucker, Tasse, Cappucino, Erwachsene
riecht gut, schmeckt gut, ist heiß, gibts in verschiedenen Varianten mit Milch, Kaffebohne, Koffein, Aufputschmittel, Herzfrequenz
wächst als Bohne, wird gemahlen, ist schwarz, macht wach
schwarz, mit Milch, Schaum , Zucker, weckt auf
wichtig, jeden morgen, trinken, gemütlcih
ist ein Lebenselexier, ist schwarz, ist aromatisch, schmeckt gut, riecht gut, macht wach, kommt aus fernen Ländern, gibt es in vielen Variationen
wachmachend, fröhlich machen, lecker, schwarz, gut riechend, heiß
trinkt man morgens, wenn man müde ist, braun, hat Aroma, man kann ihn riechen, frisch aufbrühen
Bohne, riecht gut, schwarz, mit Milch & Zucker, Latte macchiato, Espresso, braun, belebt
braun, Koffein, Zeitung
schwarz, zuckersüß, bitter
warm, bitter, schwarz, macht wach

**KÖNIGIN**
sitzt auf dem Thron, befiehlt, ist schön, ist grausam
England, Schweden, Oberhaupt, hatte mal was zu sagen, reich
intelligent, Führungsnatur, hat ein Zepter, stolz, hat Durchsetzungsvermögen, trägt eine Krone und wallende Kleider, ist sehr präsent, hat wenig Privatsphäre

Frau des Königd, schön, Krone, mächtig, schöne Gewänder, Thron, regiert mit Mann, EII, Adel
groß, Krone, dünn, regiert, abgewandt
Biene Maja
weiblicher König, einsam, mächtig, veraltet
trägt eine Krone, hat einen Thron, hat ein Königreich, ist reich, ist weiblich, hat ein Schloss, arbeitet nicht, hat Personal
kommt im Kartenspiel vor, trägt eine Krone, ist QEII, ist machtvoll
Krone, Prinzessin, Märchen, König
ist eine Schachfigur, trägt eine Krone, regiert
zickig
gibt es beim Schach, hat Krone, sitzt auf dem Thron, ist weiblich, hat roten Mantel
märchenhaft, trägt eine Krone, lebt im Schloss
vornehm, trägt Verantwortung, elegant, muss sich an die Etikette halten

**KÖNIG**
königlich, Krone, Hermelinmantel, Thron
herrschaftlich, undemokratisch, inzestuös
Märchen, Prinzessin, Königin, Schachmatt, Krone, Ludwig 14, Schloss
regiert ein Land, adelig, wohnt in Palast, wird gekrönt, hat sein amt durch Geburt, höchster Adel
regiert ein Land, hat eine Krone, sitzt auf einem Thron
Krone, Schloss, Reichtum, Gefolge
gibt es hier nicht mehr, Monarch, Thairland, Regierung, Krone, Thron, Kriege, Mittelalter
trägt im Märchen eine Krone und schöne Kleidung, wohnt im Schloss oder Palast, hat viel Geld, kaum noch Macht, eröffnet in England das Parlament, fährt mit eine Kutsche
Herrscher, gütig, mächtig, einflussreich, trägt Krone, hat ein Reich
Krone, hat ein Reich, regiert, oft böse, machthungrig, reich
Herrscher, Kartenspiel, der Oberste/Erste, Landesführer, Königin, hat Macht, entscheidet, parlamentarische Monarchie
hat eine Krone aus Gold, herrscht, sitzt auf einem Thron
reich, dick, verrückt
hat Krone, roten Mantel, Zepter

**KATZE**
ist nachts immer schwarz, ist unabhänging, ist eigensinnig, unberechenbar, will verwöhnt werden
klein, süß, haarig, kratzt gerne
kann springen, scheisst in fremde Gärten, hat verschiedene Farben, schnurrt, lässt sich gerne kraulen, jagt Vögel und Mäuse, kann Kindersatz sein
Tier, weiches Fell, krazt, beißt, schnurrt, ist den ganzen Tag faul, trinkt Milch, hat viele kleine Kätzchen
schnurrt, ist da, isst, Zufriedenheit
schlau, eigenwillig, schnurren, liebevoll, Haustier
kein Hund, selbständig, nachtaktiv, kann im Dunkeln sehen
ist klein, hat Fell, ist weich, fängt Mäuse, ist sauber, ist unabhänging, miaut
hat vier Beine, und einen Schwanz, wird gerne gestreichelt, miaut, hat Krallen, macht einen Buckel, während sie sich an Dingen reibt
fauchen, Tier, Maus, schnurren, Krallen
ist weich, ist flink, ist vorsichtig, ist verschmust
schnurrt, klettert, flauschig, nachtaktiv

ist gemütlich, schläft viel, dressiert Menschen, ist ein stolzes Tier, schnurrt, ist weich, ist biegsam
jagt Mäuse, ist sanft, hat 7 Leben, kann schnurren, eigenwillig
ist elegant, verschmust, hat Fell, hat einen Schwanz, Schnurrhaare, Krallen, schnurrt, buckelt sich

**HUND**
bellt, ist anhänglich, wedelt mit Schwanz
zeitraubend, laut, sabbernd
bellt, gassi, bewacht, Fell, kuschelig, verschiedene Rassen
bellt, muss Gassi gehen, mag spielen, ist domestiziert, ist des Menschens bester Freund, kann trainiert werden, 4 Beine, Schnauze,
bellt, hat Fell, vier Beine, Ohren, einen Schwanz
bellt, flauschig, guter Freund, Leine, Spaziergang
treuer Gefährte, bester Freund, gehorsam, anhänglich, loyal, treu, Tier, Disziplin, liebevoll
ist treu, hat schöne traurige Augen, ist schön, geht auf die Jagd, passt auf die Menschen auf, ist Begleiter, hilft behinderten Menschen, bellt, hat Fell
bellt, beißt, ist treu, ist guter Freund, kann man streicheln, hat einen Schwanz, hat liebevolle Augen
bester Freund, tappst hinterher, folgsam, unterordnend
bellt, bewacht, schnüffelt, Gassi gehen, Erziehung, Hundeschule, ist ein Rudeltier, Herrchen/ Frauchen, Hundekuchen
ist laut, bellt, hat einen Hundeblick, stinkt
stinkt, wuselig, bellen
bellt, wedelt mit dem Schwanz, hat dickes Fell

**ANGESTELLTER**
arbeitet, bekommt Lohn, hat Ferien, ist Mitglied einer Gewerkschaft, hat Pausen
unselbständig, mit oder ohne Führungserfahrung
fleissig, arbeitet im Büro
Arbeiter, Proletariat, festere Arbeitszeiten, führt Befehle aus, hat festen Lohn
arbeitet, hilft, hofft, bekommt
Normalverdienender
kriegt einen Lohn, kein Beamter, arbeitet, Steuerzahler, Gegenzug zum Arbeitgeber
arbeitet, verdient Geld, hat einen Job, kriegt Mittagspause, schreibt, telefoniert
hat einen Chef, ist Untergebener, ist nicht herausragend
Ausbeutung, Büro
ist gestresst, ist überarbeitet, trägt einen Anzug, muss viel arbeiten
unausgeschlafen
hat oft Alltagstrott, bekommt regelmäßig Geld, trinkt Kaffee, manche sind engagiert, manche sind faul, Mensch
arbeitet, untergeben, vertraglich gebunden
Untergebener, arbeitet für jemanden

**CHEF**
herrisch, laut, autoritär, männlich
fordernd, verantwortlich, respektperson
Autorität, Kündigung, Berwerbung, Führungsstelle, anti-autoritär, Betriebsleiter
leitet und verteilt Arbeit, trägt Verantwortung für Mitarbeiter
bestimmt, trägt Verantwortung, erteilt Aufträge
laut, energisch, fordernd, unbeliebt

Autoritätsperson, nicht immer gut/nett/freundlich, Vorgesetzter, Herrscher, reich, rechthaberisch, arbeiten , schuften

hat das Sagen, zahlt dein Gehalt, stellt dich ein oder aus, fällt Entscheidungen, hat Verantwortung, geht auf Geschäftsreisen, hält meetings ab

hat eine Macht-position inne, muss Entscheidungen treffen, wird häufig nicht gern gesehen, ist nicht immer beneidenswert, eine Respektsperson

hat Untergebene, ist oft laut, bestimmend, gibt Befehle, trägt Verantwortung

Vorgesetzter, über mir, Gehaltserhöhung, selbständig sein, Mitarbeiter führen und beurteilen, berichten and den Chef, Aufsicht, Führungspersönlichkeit

manchmal widersprüchlich, sorgsam, launisch

blöd, reich, langsam

Geschäftsführer, Koch, trägt Anzug und Hemd

**HAND**

hat fünf Finger, kann mit anfassen, wird dargeboten, ist das ultimative Werkzeug

5 Finger, nützlich, wichtigstes Arbeitsinstrument

kann kühl oder warm sein, ist wichtig um im Alltag funktionieren zu können

fünf Finger, sehr hilfreich, Hände schütteln, arbeiten, tippen, Teil des Körpers, streicheln, schlagen, essen, schreiben, heben, winken

reicht, Handschuhe, warm, Stark, kann

groß, klein, Handlesungen

immer im Zweierpack, hat fünf Finger, ist nürtzlich, wird täglich gebraucht

Körperteil, hat 5 Finger, ist nützlich, kann winken

hat 5 Finger, hat viele Linien, greift, ist hautfarben, hat Fingernägel dran, stellt Kontakt zur Außenwelt dar

Finger, Körper, waschen

ist klein, ist groß, ist breit, schmal, kalt, warm

essentiell

hat 5 Finger, der Mensch at 2 davon, ist Werkzeug

schafft etwas, kann anpacken, hat 5 Finger

hat Finger, Nägel, einen Daumen, Handlinien, verschiedene Formen und Temperaturen, wird bei Nervosität feucht, Sammelpunkt vieler Nerven, gut um Krankheiten zu diagnostizieren

**HANDSCHUH**

wärmt, aus Leder oder Wolle, man findet immer nur einen

warm, aus Leder, winterlich

hält warm, man verliert immer einen, Ohrfeigen, Winter

warm, hält Hände warm

wärmt, schützt Hände vor Kälte, und Schmutz

warm, Winter, nass

spendet Wärme, schützt bei Sport, manchmal notwendig, Leder, Daunen, Fingerling, Fäustling

hält warm, ist aus Leder oder Wolle, Fingerhandschuh oder Fäustling, schützt die Haut, ist praktisch

hat fünf Finger, wärmt, ist schick, verliert man leicht, ist aus Wolle, wichtig zum Fahrrad fahren

sind oft rot, können Fausthandschuhe sein, oft gestrickt, warm

schützt im Winter vor Kälte, schützt vor Schmutz und Verletzungen, gestrickt oder aus Fleece, Gummihandschuhe, Fäustlinge, Fingerhandschuhe

wärmt, gefüttert, skifahren, nicht nass

geht schnell verloren, warm

trägt man wenn's kalt ist, ist schwarz und aus Leder, kann auch aus Wolle sein

**MUTTER**
sorgt sich, kümmert sich, behält den überblick, hat Kinder, hat einen Mann
ist hoch zu achten, hat Kinder, ist oft zu Hause
sorgt sich um die Familie, hat ein offenes Ohr, hat ein offenes Ohr, hat viele Ratschläge, ist immer für einen da, unterstützt die Kinder, kocht und macht Hausarbeit und managt die Familie
hat mir das Leben geschenkt, Liebe, erzieht, schimpft, lacht, hilft, ist immer für ihr Kind da
lieb, gibt, bekommt, ist da, Beschützerin, Frau des Vaters
Bernadette, erziehen, liebevoll
jeder hat eine, sollte mütterlich sein, ist Erzieherin, Mitbegründerin einer Familie
ist lieb, passt auf, ist weiblich, ist groß, ist fürsorglich, kocht und backt
ist warm, gibt Geborgenheit, ist immer da, sorgt sich, lacht
Vater, helfen, Geburt
ist lieb, fürsorglich, geduldig, schimpft
menschlich, fürsorglich, verantwortungsbewusst
ist Ernährerin, hat Brüste, ist weiblich, ist weich
hat geboren, fürsorglich, weiblich, Erde
warm, liebevoll, fürsorglich

**VATER**
autoritär, männlich, abwesend
fürsorglich, gefährlich, streng
mindestens ein Kind, fürsorglich, Oberhaupt der Familie, ist ein Mann
ist ein Mann, zeugt Kinder, liest vor, baut und spielt mir Kindern, ernährt die Familie mit
Oberhaupt der Familie, männlich
ist dein Erzeuger
ist häufig Erzeuger, streng, liebevoll, besorgt
Vorbildfunktion, Erziehungsfunktion, liebevoll, beschützend
Mutter, Kind, ist Vorbild, wird respekt gezollt, beschützt Kinder, sorgt für die Familie
Gegner, Beschützer, Familie
nett, gemütlich
Mann, alt, Fußball, Beschützer

**TASSE**
enthält Tee, Kaffee, Milch, muss bestimmte Ausmaße haben, hat einen Hänkel
Behälter für Kaffee
gibt es in verschiedenen Formen und Größen, ist hart und stabil, ist ein Hilfsmittel um Tee zu trinken
Gefäß zum trinken, kann groß und klein sein, handlich, Kaffee, Tee, bunt, einfach, Motiv, Spruch
fasst, gerne, groß, leer, voll
Beschriftungen, praktisch
ist zum trinken da, beinhaltet Flüssigkeit, Tee, Kaffee
ist ein Gegenstand, ist ein Behälter, ist klein, hat einen Griff, ist oft bemalt, ist oft aus Porzellan
hat einen Henkel, ist oft zylinderförmig, wird mit Tee oder Kaffee befüllt, wird mit beiden Händen gegriffen
trinken, Henkel, Geschirr
ist zerbrechlich, steht im Schrank, ist groß oder klein, ist aus Porzellan
einfach zu greifen
ist Gefäß, hat Henkel
Zerbrechlich, hat einen Henkel, ist nützlich

besteht oft aus Porzellan, aus ihr wird Kaffee oder Tee getrunken, ist oft bemustert, hat Sprüche oder Namen drauf

**BECHER**

alt mit Teeflecken, hat einen Sprung, hält warm, spendet Vertrauen
zivilisatorisch, aus Glas, schlicht
trinken, Plastik, Spülmaschine, Pfand
daraus kann mann trinken, nicht wertvoll, oft aus Plastik, hat keinen Hänkel
Trinkgefäß, gibt es aus verschiedenen Materialien, hat einen Henkel, hat einen Boden, ist oben offen
Wasser, hat Henkel, ist warm oder kalt
trinken, Gefäß, Plastik, Joghurt, Joghurtbecher
ist ein Trinkgefäß, gibt es aus Glas oder Porzellan oder Plastik, ist für Wasser oder Saft
nützlich, trinkt man draus, wäscht man Pinsel drin aus, aus Glas, aus Plastik, kann zum Zähneputzen verwendet werden
rund, kann aus Plastik sein, kommt Wasser rein, man kann daraus trinken
ist halb leer, halb voll, Gefäß zum trinken, Eisbecher, Joghurtbecher
zum trinken, aus plastik, praktisch fürs Camping
warm, mit Kaffee gefüllt
aus Metall oder Glas, kann auch aus Porzellan oder Steingut sein, man trinkt Wein draus

**ZWERG**

ist klein, hat einen Bart, trägt keine Zipfelmütze, baut Erz und Mineralien ab, kennt sich im Schmieden von Waffen aus, bekämpft Drachen
klein, vorlaut
kommt im Märchen vor, ist klein, frech, schlau
klein, Märchen, Schneewittchen, Garten, Dekorativ, bieder, kitschig, rote Mütze
klein, Gnom, sieht alles, Garten, Tomte
klein, Garten, Zipfelmütze, Spaten
klein, kleinwüchsig, hat mythological, Garten
ist ein Mensch, ist klein
hat eine Zipfelmütze auf, hat manchmal eine große Nase, ist klein, ist robust und bodenständig
klein, Nase, Märchen, Zipfelmütze
klein, mürrisch, trägt eine Mütze, hat einen Bart, ist alt, lebt im Wald
klein, listig, kitschig
ist klein, Wesen aus Märchen, hat Bart, hat Mütze
klein, trägt eine Mütze, arbeitet hart
klein, hat eine Zipfelmütze, manchmal einen Bart, gibt es als Gartendekoration, trägt oft Latzhosen

**RIESE**

häßlich, groß, Mythologie, gefährlich
unmenschlich, gefährlich, einsam
Märchen, Gebrige, Zwerg, groß, übernatürlich
ist sehr groß, Märchenwesen oder nur großer Mensch
ist groß, ist eine Märchengestalt, ist oft böse
kommt in Märchen vor, David und Goliath
groß, Hüne, Märchengestalt, Basketball, Sport, erfolgreich
groß, stark, im Märchen meistens böse und plump
Gegenzatz zu Zwerg, groß, mächtig, spielt in Märchen/Sagen mit, gelten als bedrohlich
ist groß, Gegensatz von Zwerg, David und Goliath

ist sehr groß, Gegenteil ist der Zwerg, Siebenmeilenstiefel, unüberwindbar, unermesslich, übergroß
groß, David gegen Goliath, stark, kräftig
groß, böse, isst viel, hat große Füße
groß, männlich, kommt im Märchen vor, ist meistens grimmig

**WIND**
bewegt Luft, geht von da nach dort, kann kräftig blasen, oder leise sein, kann zum Sturm werden
kalt oder warm
lässt einem die Augen tränen, trägt Segelschiffe aufs weite Meer, lässt einem Gänsehaut kriegen
Wetter, Sturm, Lufthauch, nicht sichtbar, stark oder schwach, Blätter wirbeln, Rückenwind
bläst, kalt, Holland, weht, Fahne, Haare im Gesicht
Brise
ist unsichtbar, kann milde oder kräftig sein, irritierend wenn mit Regen gekoppelt, kann energiespendend sein
ist unsichtbar, kalt, erfrischend, stark, weht, bewegt
fühlt sich gut auf der Haut an, unsichtbar, schmeichelnd, zerstörerisch, unberechenbar, man kann darauf fliegen, ist kraftvoll oder sanft
Hauch, Kälte, Sturm
ist erfrischend, ist kalt, ist stark, wirbelt Dreck auf, weht
kalt, angenehm
ist stürmisch, ist sanft, pustet, weht, verweht Dinge, ist stark
kalt, brausend, nordischen Phänomen, rein
erfrischend, nervig, stark oder schwach, kann Ohrenschmerzen verursachen

**REGEN**
naß, schön, viel Romantik
nass, grau, frische Luft
nass, gut für Pflanzen, unangenehm, Regenschirm, graues Wetter
gibt es viel in England, Wasser aus den Wolken, verschiedene Formen, macht Land fruchtbar, Wetterbedingung
fällt in Tropfen auf die Erde, entsteht in den Wolken, ist nass
Tropfen, schwarze Wolken, Regenschirm
frustig, lauer Sommer-regen schön, Regenbogen, kalt oder warm, leben spendend, notwendig für Menschen, Tier und Natur
kommt aus dicken Wolken gefallen, ist nass, ist im Sommer manchmal schön, ist wichtig für die Natur, ist oft kalt und ungemütlich, gibt es als Nieselregen oder Sturzbachregen
erfrischend, naß, kalt, angenehm, macht Geräusche
ist naß, kalt, trist
ist nass, kühlt, wässert Pflanzen, Wasser, Tropfen, Eisregen, Regenschirme schützen vor Regen, Regenmantel
nass, Wetter, niedriger Luftdruck
nass, kalt
nass, kalt, er macht Pfützen, man braucht Regenschirm oder Regenjacke

**MANN**
denkt praktisch, ist Ziel orientiert, ist stark, bringt Geld ins Haus
Krönung der Schöpfung
gross, stark, intelligent, schön beschäftigt

Gegenteil von Frau, stark, groß, andere Geschlechtsorgane, Beschützerinstinkt, Autos, Fußball, sorgt sich um Familie
Frau, groß, will, macht, sagt
groß, stark, muskulös, Arbeiter
keine Frau, männlich, stereotypisch kräftig, Lohnverdiener, Jäger, hat einen Preis
hat einen Bart, baut
ist kastenförmig, hat eine tiefe Stimme, hat Kraft, ist aktiv, redet oft nicht gern über Gefühle, ist praktisch
groß, blöd, Frau
ist groß, ist stark, trinkt Bier
menschlich, normal
ist Mensch, hat Bartwuchs, hat keine Brüste
groß, stark, hat einen Bart
solte stark mutig und verantwortungsbewusst sein, hat einen Penis, starke Brust- Rücken- Ohrenhaare, leidet oft unter Haarausfall, Testosteron

**FRAU**
weiblich
organisationstalent, immer auf Trab, weiblich
Eva, Liebe, Schwangerschaft, Geburt, Mutter, schenkt Leben, Weiblichkeit, Mann
weiblich, Person, erwachsen, oft Ehefrau
hat Brüste, bekommt Kinder, arbeitet, wäscht, kocht, putzt
weiblich, Brüste, Mutter, analysiert gern
feminin, zierloch, weiblich, liebevoll, versorgend, Familie zusammenhalten, Haushalt werfen, Kinder gebähren
ist weiblich, schön, trägt oft Kleider, liebt schöne Schuhe, ist meistens kleiner als der Mann
ist eine Geschlechtszuordnung, hat bestimmte Geschlechtsorgane, wird in Zusammenhang Zuneigung gebracht, weiche Formen, hohe Stimme, kann Kinder gebären
feminin, weiblich, kann Kinder bekommen
Mann, eitel, schön, schwaches Geschlecht, hat Kurven
weich, schön, laut , bietet Kontraste
schön, anstrengend
lange Haare, trägt auch Röcke oder Kleider, mag Schuhe

**BERG**
ist hoch, ist schneebedeckt, ist aus Stein, will bezwungen werden, steht über allem
hoch, anstrengend zu besteigen, hat evtl Seilbahn, zeigt schöne Aussicht, verbindet Heimatgefüh
kann man besteigen, ist mächtig, steht einfach da, bildet tolle Landschaft
Gebirge, hoch, Gestein, steil, erklimmen, schöne Aussicht, anstrengend, Alpen
hoch, Tal, Schweiz, Arbeit, Mühsahl
Schnee, Skifahren, Bergbahnen, Gipfel, Musik, wandern, Ablenkung, Natur
gross, idealerweise mit Felsen, spärliche Vegetation, schwer zu erklimmen, sollte ewigen Schnee auf dem Gipfel haben
ist hoch, massiv, ist schneebedeckt, steil
groß, ist aus Stein, bietet einen Ausblick auf die Welt an, kann bestiegen werden, hat Höhlen und Spalten
hoch, Gebirge, steigen, wandern, skifahren, Alpen
ist mit Gras bewachsen, ist felsig und steil, ist groß, beherbergt Gemsen und Murmeltiere
groß
ist hoch, Fels, man kann darauf wandern, klettern

felsig, hoch, kalt, eisig, ein natürliches Hinderniss, eine Herausforderung
ist steil, mit Bäumen oder Gras bewachsen, hat einen Gipfel auf dem Schnee und Eis sein können

**TAL**
neblig, geheimnissvoll, mysteriös, unten
Natur, Abgeschiedenheit, gefährlich
Berg, dunkel, Flüsse
ist am Fuß, zwischen Bergen
ist unten am Berg
zwischen hohen Bergen, wandern, Dorf
wo ein Tal da auch Berge, wandern, Flußtal, geschlängelt, grün, Natur, tief, schroff
liegt zwischen 2 Bergen, liegt im Schatten der Berge, ist grün, hat oft einen Fluss
unterhalb eines Berges, hat einen Fluss oder Bach, kann man Spazierengehen, ist einen Landschaft
bei Bergen, aushölung, liegt unten, dort sind oft Siedlungen
liegt zwischen Bergen, Berg und Tal, tiefes Tal, geht nach unten, am Fuße des Berges
meist schön, hat einen Fluss und Wiesen, ist umrandet von Bergen
tief, schön, dunkel
liegt zwischen Bergen, oft mit Bach oder Fluss, Kühe stehen dort manchmal auf der Weide

**LEHRER**
lehrt, ist ein Vorbild, hat es schwer, sollte allwissend sein
Soziologe, könnte mit Strenge etwas erreichen, ist soft
weiss immer alles besser, ist streng
Erzieher, unterrichtet Schüler, Schule, Klassenzimmer, Hausaufgaben, Wissen, lange Ferien, korrigieren
Arbeit, gut, Schüler, anstrengend, hilft, korrigiert
Lehrender, vermittelt Wissen, generiert Interesse, weiss mehr als die Lernenden
ist in der Schule, hat einen Aktenkoffer, korrigiert Tests, ist streng, ist nett
bringt Dinge bei, verbietet, gibt Noten, ist seltsam, ist von sich überzeugt
Schule, Tafel, Noten, autoritär
trägt eine Brille, unterrichtet Kinder, schreibt an die Tafel, gibt Noten, tadelt Schüler
verantwortungs-bewusst, Vorbild, unterbewertet
ist gütig, ist gerecht, hat Wissen, soll es vermitteln, hat große Tasche, hat Rotstift
wissend, spricht viel und gerne
ist klug, belesen, hat Vorbildfunktion, kann streng sein, gibt Hausaufgaben, unterrichtet an Schulen, schreibt mit Kreide oder Filzstift and die Tafel

**SCHÜLER**
jung, laut, zappelig, aufmerksam, unaufmerksam, kindisch
pubertär, rechthaberisch, faul
muss lernen, Lehrer, Prüfung
geht zur Schule, lernt bekommt Noten, muss Hausaufgaben machen
meistens ein junger Mensch, lernt, erlernt Fähigkeiten
Schule, Schulaufgaben, Schulranzen, Tafel, Lehrer
jung, muss lernen, wird erwachsen, Teenager, ausprobieren, experimentieren, Leben lernen, Jugendliche ohne Werte
bekommt eine Schultüte, ist oft noch ganz klein, muss noch viel lernen, spielt während der Pausen, muss Hausaufgaben machen, hat Spaß am Lernen, schreibt Klausuren und Prüfungen, lernt für's Leben

geht zur Schule, soll lernen, hat einen Lehrer, mag die Schule nicht
lernen, sind zusammen, Klassenverband
geht zur Schule, muss etwas lernen, Schulpflicht, wird vom Lehrer unterrichtet
soll lernen, ist jung, hat einen Tornister, hat noch alles vor sich
jung, laut, anstrengend
sind laut, haben Schulranzen, ärgern Lehrer, müssen lernen, gehen in eine Klasse

## **Appendix 11** Translations of German data

### Chapter 5.1.1

| | |
|---|---|
| traurig:unglücklich | sad:unhappy |
| neu:alt | new:old |
| oben:unten | up:down, top:bottom |
| groß:klein | big:small, short:tall |
| packen:auspacken | pack:unpack |
| lernen:verlernen | learn:unlearn (forget) |
| Zwerg:Riese | dwarf:giant |
| Tod:Leben | death:life |
| Mann:Frau | man:woman and husband:wife |
| Berg:Tal | mountain:valley |
| nehmen:geben | take:give |
| ausziehen:anziehen | undress:dress |
| erlauben:verbieten | allow:forbid |
| kaufen:verkaufen | buy:sell |
| fragen:antworten | ask:answer |

### Chapter 5.1.2

| | |
|---|---|
| clever:helle | clever:smart (literally ‘light’) |
| Buch:Seite | book:page |
| blau:orange | blue:orange |
| Krone:König | crown:king |

### Figure 5.4

| **MUTTER** | |
|---|---|
| ist fürsorglich | is caring |
| ist warm/lieb | is warm/nice |
| hat geboren | has given birth |
| erzieht | educates/brings up |
| ist weiblich | is female |
| ist immer da | is always there |
| hilft/unterstützt | helps/supports |
| schimpft | tells you off |
| lacht | laughs |
| kocht | cooks |

**Figure 5.5**

| VATER | |
|---|---|
| ist ein Mann/männlich | is a man/male |
| ist ein Erzeuger | is a genitor |
| ist fürsorglich/liebevoll | is caring |
| beschützt | protects |
| ist Oberhaupt der Familie | is the head of the family |
| ist ein Vorbild | is a role model |
| ist streng | is strict |
| sorgt für Familie | provides for the family |

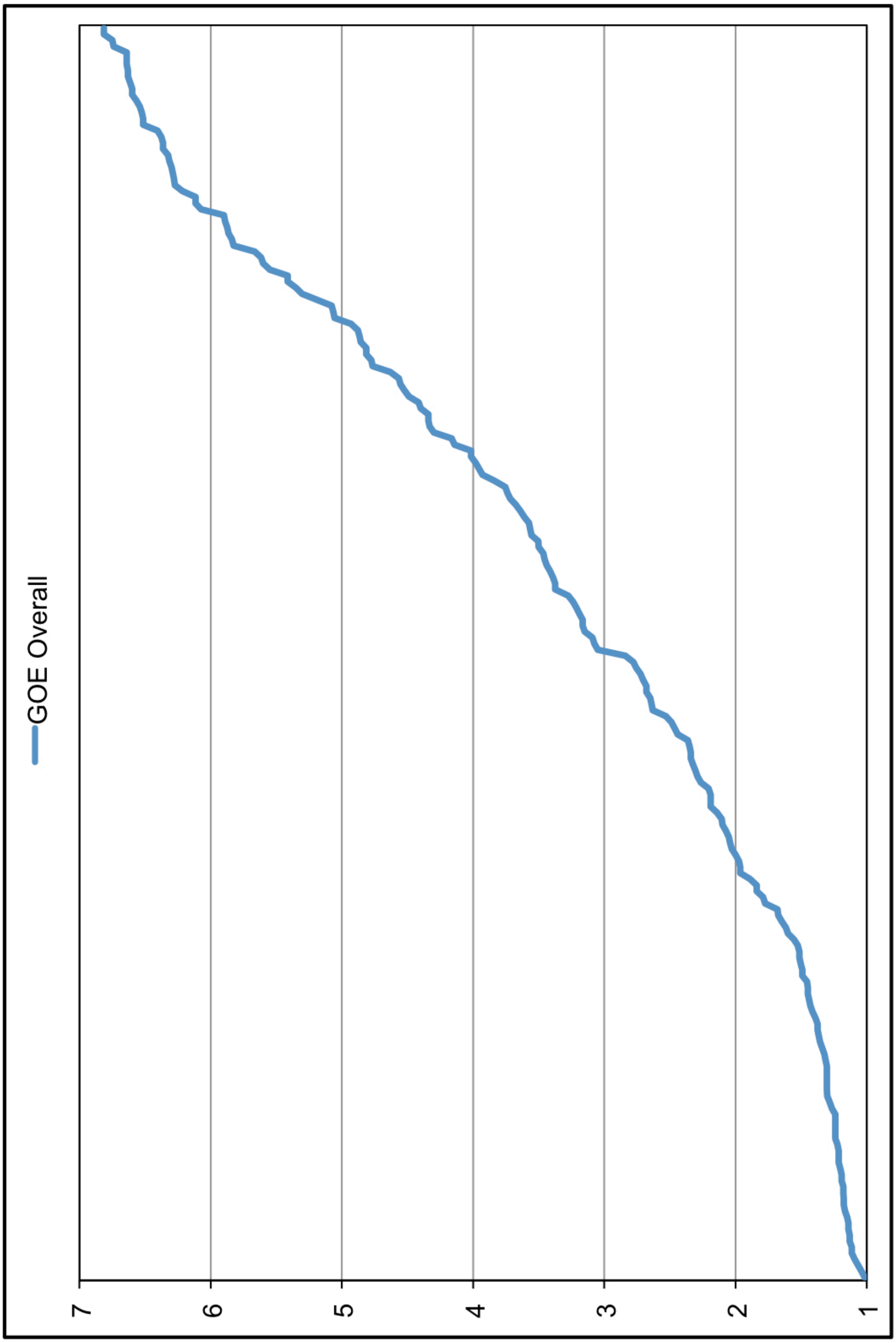
**Figure 5.6**

| FRAU | |
|---|---|
| weiblich | is female |
| schön | beautiful |
| kann Kinder bekommen | can bear children |
| hat Brüste | has breasts |
| liebevoll | is caring |
| feminin | feminine |
| weich | soft |
| trägt Kleider/Röcke | wears dresses/skirts |
| zierlich/kleiner als Mann | is dainty, smaller than a man |
| Mutter | mother |
| macht Haushalt | does chores |

**Figure 5.7**

| MANN | |
|---|---|
| stark | strong |
| groß | tall |
| Frau/Gegenteil von Frau | woman/opposite of woman |
| hat einen Bart | has a beard |
| ist ein Mensch/menschlich | is (a) human |
| verdient Geld | earns money |
| praktisch | is practical |
| hat einen Penis | has a penis |
| trinkt Bier | drinks beer |
| männlich | male |

**Appendix 12** Chart of all English pairs sorted by GOE-rating



**Appendix 13** Pairs used for Figure 6.3

| ANTONYMS | UNRELATED |
|---|---|
| LARGE-SMALL | EAGER-CLEAR |
| LEFT-RIGHT | CLOCK-SONG |
| ORGANISED-DISORGANISED | COMPETENT-INCONSISTENT |
| LITTLE-LARGE | CONTEXT-MACHINE |
| OLD-NEW | DEAD-SAFE |
| PALE-LIGHT | IMMATURE-IDEAL |
| SAD-HAPPY | INFERIOR-CONSERVATIVE |
| SHORT-TALL | KEEN-NICE |
| SMALL-BIG | FISHWIFE-CRITIC |
| SWEET-SOUR | EMPTY-IMPARTIAL |
| WRONG-RIGHT | ENGINEER-RECEIVER |
| TINY-BIG | REPAIR-INSULT |
| UNEVEN-EVEN | PRACTICAL-FIDGETY |
| UNINTERESTED-INTERESTED | NEUROTIC-IMPATIENT |
| UNORGANISED-ORGANISED | NEED-PUT |
| HAPPY-UNHAPPY | MODERN-GRAY |
| GOOD-BAD | ORDER-MEDIATOR |
| EVEN-ODD | SOUTH-WINTER |
| EVIL-GOOD | STEAM-ZONE |
| FEMALE-MALE | SUBJECT-TOWN |
| FEMININE-MASCULINE | THIN-FIRM |
| EASY-HARD | UGLY-WIDE |
| HARD-SOFT | UNPOPULAR-INFERIOR |
| HEALTHY-SICK | WIDE-HEAVY |
| HEALTHY-UNHEALTHY | WILD-EERIE |
| INTERESTED-DISINTERESTED | WINDOW-SERVANT |
| HOT-COLD | FUNNY-WILD |
| HUGE-SMALL | FREE-LATE |
| HUGE-TINY | FULL-BROAD |
| YOUNG-OLD | IDEAL-ILL |
| COOL-HOT | CALM-MUDDY |
| COOL-WARM | BELIEF-TALENT |
| BIG-LITTLE | |
| BITTER-SWEET | |

| | |
|---|---|
| BOILING-FREEZING | **RELATED** |
| **SYNONYMS** | TRIBE-VILLAGE |
| BROAD-WIDE | PENNY-DOLLAR |
| BULLET-MISSILE | QUIET-EMPTY |
| CLEAN-FRESH | SNOW-ICE |
| CLEAR-BRIGHT | GUESS-KNOW |
| LOOK-SEE | FOREST-PLANT |
| SAFE-WELL | FOREIGN-CULTURE |
| MUDDY-UNTIDY | FIRM-FRESH |
| HOG-PIG | HEAVY-FULL |
| KIND-NICE | IDLE-FREE |
| ANGRY-MAD | HERE-SOON |
| ANGRY-UNSTABLE | DRIVER-CAPTAIN |

## **Appendix 14** Index of measures used in analyses

### GOE-RATING

The goodness-of-exemplar rating is the mean of the ratings given by all participants for a given word pair. The ratings range from **1 (excellent)** to **7 (very poor)** and are given to the 3rd decimal place wherever necessary. Thus, the lower the score, the higher the antonymic strength of a pair.

### T-SCORE

The t-score is the measure used to indicate whether frequency of co-occurrence in the BNC (COSMAS for German examples) is likely to be due to chance rather than due to any linguistic effects. A t-score below 2.0 is very low and does not allow any firm conclusions to be drawn. Any co-occurrence with a t-score above 2.0 is unlikely to be due to chance. The higher the t-score, the more likely it is that there is a linguistic cause for the co-occurrence.

The t-score is a measure which is sensitive to sample size and thus the t-scores in the German data are higher since COSMAS is, at 2 billion words, a much larger corpus than the BNC. This, however, does not render the t-score measures superfluous in comparing the data, since it is the differences **within** the languages that are being compared rather than an in-between language comparison of the t-score.

### EAT SCORE

The Edinburgh Word Association Thesaurus is a good indicator of associative strength, despite the fairly small sample size of 100 responses per word. The data shows how many participants out of the total 100 responded with the second member of a word pair when presented with the first. Anything above 0.30 can be considered extremely high.

### MEAN RT

Mean RT is the mean reaction time of all correct responses to a certain word pair. Null responses as well as wrong responses have been taken out to avoid any bias in the data. In all ANOVAs, the least square mean is reported since this is a more accurate indicator than the simple mean.