

# Contents

|                               |      |
|-------------------------------|------|
| The author and series editors | xv   |
| Foreword                      | xvii |
| Introduction                  | 1    |
| <b>1 Corpus creation</b>      |      |
| Introduction                  | 13   |
| Who should design a corpus?   | 13   |
| A general corpus              | 13   |
| Outline of corpus creation    | 14   |
| Electronic form               | 14   |
| Permissions                   | 15   |
| Design                        | 15   |
| Spoken and written language   | 15   |
| Quasi-speech                  | 16   |
| Formal and literary language  | 16   |
| Typicality                    | 17   |
| Design criteria               | 17   |
| Period                        | 18   |
| Overall size                  | 18   |
| Sample size                   | 19   |
| Whole documents               | 19   |
| Minimal criteria              | 20   |
| Provisional corpus            | 20   |
| Processing                    | 20   |
| Clean-text policy             | 21   |
| Basic provision               | 22   |
| Database                      | 22   |
| Maintenance                   | 22   |
| Different kinds of corpora    | 23   |
| Sample corpora                | 23   |
| Monitor corpora               | 24   |
| Features of a monitor corpus  | 25   |

## 2 Basic text processing

|                                 |    |
|---------------------------------|----|
| Introduction                    | 27 |
| Input                           | 28 |
| Words and word-forms            | 28 |
| Text and vocabulary             | 29 |
| Frequency list—first occurrence | 30 |
| Frequency list—alphabetical     | 31 |
| Frequency list—frequency order  | 31 |
| Word frequency profiles         | 32 |
| Concordances                    | 32 |
| KWIC (Key Word in Context)      | 32 |
| Longer environments             | 33 |
| Ordering within concordances    | 33 |
| Concordance processing          | 34 |
| Text analysis statistics        | 34 |
| Selective information           | 35 |
| Intermediate categories         | 35 |
| New approaches                  | 36 |

## 3 The evidence of usage

|                                  |    |
|----------------------------------|----|
| Introduction                     | 37 |
| Existing descriptions            | 37 |
| Native-speaker introspections    | 39 |
| Language in use                  | 39 |
| Word-forms and lemmas            | 41 |
| Concordances                     | 42 |
| Concordance evidence: an example | 44 |
| <i>Sense 1: to refuse</i>        | 47 |
| <i>Other senses</i>              | 49 |

## 4 Sense and structure in lexis

|   |    |
|---|----|
| Introduction                              | 53 |
| Evidence: main senses                     | 53 |
| Evidence: minor senses                    | 55 |
| Counter-examples: general                 | 56 |
| Counter-examples: first sense             | 57 |
| <i>Yield</i> with an object               | 58 |
| <i>Yielding</i> with an object            | 59 |
| <i>Yielded</i> with an object             | 60 |
| Descriptive and prescriptive study        | 60 |
| Counter-examples: second and third senses | 61 |
| <i>Yield</i> as transitive verb           | 61 |
| Doubtful cases                            | 62 |
| First minor sense                         | 63 |
| Conclusion                                | 65 |

## 5 Words and phrases

|                                       |    |
|---------------------------------------|----|
| Introduction                          | 67 |
| Phrasal verbs                         | 67 |
| Some numerical facts                  | 68 |
| Combinations of <i>set</i> + particle | 69 |
| The combination <i>set in</i>         | 70 |
| Nouns                                 | 72 |
| Verbs                                 | 72 |
| <i>Sense (i)</i>                      | 72 |
| <i>Sense (ii)</i>                     | 72 |
| <i>Minor senses</i>                   | 73 |
| <i>Sundry idioms</i>                  | 73 |
| <i>Set in</i> as a phrasal verb       | 73 |
| <i>Word-forms</i>                     | 74 |
| <i>Subjects</i>                       | 74 |
| <i>A dictionary entry</i>             | 75 |
| Other phrasal verbs with <i>set</i>   | 75 |
| Conclusion                            | 78 |

## 6 The meeting of lexis and grammar

|  |    |
|--|----|
| Introduction                               | 81 |
| What is said about <i>of</i>               | 81 |
| A corpus view of <i>of</i>                 | 82 |
| Frequency                                  | 84 |
| <i>Of</i> outside nominal groups           | 85 |
| <i>Of</i> in nominal groups                | 85 |
| Conventional measures                      | 85 |
| Less conventional measures                 | 86 |
| The status of headword                     | 86 |
| Focus nouns                                | 87 |
| Focus on a part                            | 87 |
| Focus on a specialized part                | 88 |
| Focus on a component, aspect, or attribute | 88 |
| Support                                    | 89 |
| Metaphor                                   | 90 |
| Double-headed nominal groups               | 90 |
| Titles                                     | 90 |
| Nominalizations                            | 91 |
| Modification of first noun (N1)            | 93 |
| Mopping up                                 | 94 |
| Superlative adjectives                     | 94 |
| Fixed phrases                              | 94 |
| Miscellaneous                              | 94 |
| Evaluation                                 | 95 |
| Analysis of examples in Table 1            | 96 |
| Non-nominal instances of <i>of</i>         | 96 |
| Nominal group                              | 96 |
| Conclusion                                 | 98 |

## 7 Evaluating instances

|                          |     |
|--------------------------|-----|
| Introduction             | 99  |
| Throw away your evidence | 99  |
| Text and language        | 102 |
| Meaning and structure    | 104 |
| Procedure                | 105 |
| Findings                 | 107 |
| Conclusion               | 108 |

## 8 Collocation

|  |     |
|--|-----|
| Introduction   | 109 |
| Two models of interpretation                         | 109 |
| The open-choice principle                            | 109 |
| The idiom principle                                  | 110 |
| <i>Evidence from long texts</i>                      | 112 |
| Collocation  | 115 |
| Collocation of <i>back</i>                           | 116 |
| Analysis of the collocational pattern of <i>back</i> | 117 |
| <i>Upward collocates: back</i>                       | 117 |
| <i>Downward collocates: back</i>                     | 118 |
| Conclusion   | 121 |

## 9 Words about words

|                              |     |
|------------------------------|-----|
| Introduction                 | 123 |
| Structure                    | 124 |
| Variation in co-text         | 126 |
| About the word itself        | 126 |
| What people mean             | 126 |
| Structure: verb explanations | 127 |
| Animate subjects             | 127 |
| Inanimate subjects           | 129 |
| Mixed subjects               | 130 |
| Operators                    | 130 |
| Summary                      | 132 |

## *Contents*

|                             |     |
|-----------------------------|-----|
| Analysis of the second part | 132 |
| First chunk                 | 132 |
| Second chunk                | 134 |
| Discussion                  | 134 |
| Applications                | 135 |
| Inferences and implications | 135 |
| Conclusion                  | 136 |
| Summing up                  | 137 |
| <br>                        |     |
| Appendix I                  | 139 |
| Appendix II                 | 150 |
| Appendix III                | 154 |
| Bibliography                | 165 |
| Glossary                    | 169 |
| Index                       | 177 |