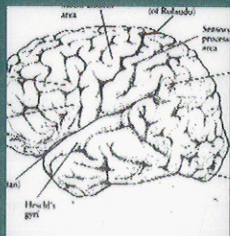
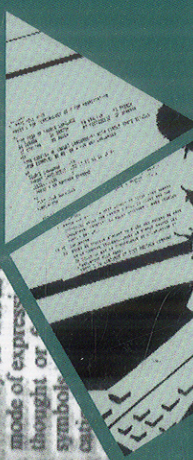


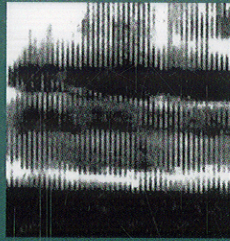


TYPE CODE OF SOURCE LANGUAGE  
DG GERMAN DK DAN  
IT ITALIAN NG DUTCH  
EG  
TYPE CODE(S) OF TARGET LANGUAGE  
(FOR EXAMPLE: DG NG)  
A  
SOURCE LANGUAGE : EN  
TARGET LANGUAGE(S) : DG  
SUBJECT CODE :



...the  
...instrument  
...language long  
...or body of words  
...mode of expression  
...thought or  
...symbols  
...can

...then just hope  
...then nig  
...to sleep on  
...me nothing  
...wa's  
...rding



...with  
...their writing  
...LEXICOGRAPHY.  
...The art or practice  
...ries.  
...EXICON  
...ary; a book  
...words  
...Th

# CORPUS LINGUISTICS

Investigating Language  
Structure and Use

Douglas Biber,  
Susan Conrad & Randi Reppen

# Contents

|  |                |
|--|----------------|
| <i>Preface</i>   | <i>Page ix</i> |
| 1 Introduction: goals and methods of the corpus-based approach | 1              |
| <b>Part I Investigating the use of language features</b>       | 19             |
| 2 Lexicography   | 21             |
| 3 Grammar  | 55             |
| 4 Lexico-grammar   | 84             |
| 5 The study of discourse characteristics                       | 106            |
| <b>Part II Investigating the characteristics of varieties</b>  | 133            |
| 6 Register variation and English for Specific Purposes         | 135            |
| 7 Language acquisition and development                         | 172            |
| 8 Historical and stylistic investigations                      | 203            |
| <b>Part III Summing up and looking ahead</b>                   | 231            |
| 9 Conclusion   | 233            |
| <b>Part IV Methodology boxes</b>                               | 243            |
| 1 Issues in corpus design                                      | 246            |
| 2 Issues in diachronic corpus design                           | 251            |

|             |   |     |
|-------------|---|-----|
| <i>viii</i> | <i>Contents</i>   |     |
| 3           | Concordancing packages versus programming for corpus analysis | 254 |
| 4           | Characteristics of tagged corpora                             | 257 |
| 5           | The process of tagging  | 261 |
| 6           | Norming frequency counts                                      | 263 |
| 7           | Statistical measures of lexical associations                  | 265 |
| 8           | The unit of analysis in corpus-based studies                  | 269 |
| 9           | Significance tests and the reporting of statistics            | 275 |
| 10          | Factor loadings and dimension scores                          | 278 |
|             | Appendix: commercially available corpora and analytical tools | 281 |
|             | <i>References</i>   | 288 |
|             | <i>Index</i>  | 297 |