Contents

Thanks and Acknowledgements 9
1. Introduction 11
1.1 Background and Purpose of this Study 11
1.2 The Theoretical and Methodological Framework 13
1.3 Outline of Content 21

2. A Corpus of Reading Comprehension Tests 23
2.1 The DIALANG Sub-corpus 24
2.2 The TOEFL Sub-corpus 26
2.3 The IELTS Sub-corpus 30
2.4 The PET Sub-corpus 32
2.5 The FCE Sub-corpus 34
2.6 The CAE Sub-corpus 36

3. Quantitative Investigation into the Complexity of the Texts of the Corpus 41
3.1 The Type/Token Ratio 42
3.2 Measuring Lexical Density and Lexical Variation 48
3.2.1 Measures of Lexical Density 49
3.2.1.1 Lexical Items vs. Grammatical Words 53
3.2.1.2 Lexical Density (Ure's method) in the Corpus 58
3.2.2 A Measure of Lexical Variation 64
3.3 Measuring Vocabulary Difficulty 71
3.3.1 Range 32 and the Underlying Research 72
3.3.2 Using Range 32 and Base Word Lists for the Analysis of the Corpus 75
3.4 Measuring Grammatical Intricacy and Halliday's method of Measuring Lexical Density 87

3.4.1 Halliday's Method of Calculating Lexical Density (LDH), Grammatical Intricacy, the Grammar of the Clause Complex, and Embedding 88
3.4.1.1 The Clause Complex 89
3.4.1.2 Embedding of Clauses 94
3.4.2 Further Investigation of the Lexical Density of the Texts and an Investigation of their Grammatical Intricacy 97
3.5 Measuring Readability 103
3.5.1 Readability Formulae 103
3.5.2 Investigating the Readability of the Corpus 106
3.6 On the Relationship between the Variables 113

4. Text Complexity, Participant Identification and Cohesion in the Corpus 123
4.1 Participant Identification in English: from Discourse Semantics to Lexicogrammar 124
4.1.1 Participant Identification in English: a Review of some Relevant Contributions 124
4.1.2 Phoricity and the Tracking System Network 129
4.1.3 The Participant Identification System Network 140
4.2 Semantics, Lexicogrammar and the Textual Metafunction: a Systemic Functional Perspective 147
4.2.1 The Semantic and the Lexicogrammatical Strata in Systemic Functional Linguistics 148
4.2.2 The Textual Metafunction: Cohesion and Thematic Structure in English 150
4.2.2.1 Cohesion in English 151
4.2.2.2 Thematic Structure 153
4.3 Participant Identification and the English Nominal Group 157
4.4 The Role of Participant Identification and of Textual Cohesion in the Corpus 164
4.4.1 Test 12 165
4.4.2 Tests 2, 25 and 11 170
4.4.2.1 Test 2 171

4.4.2.2 Test 25 174
4.4.2.3 Test 11 181
4.4.3 Test 14 184
4.4.4 Test 22 189
4.4.5 Tests 15, 3 and 13 193
4.4.5.1 Test 15 194
4.4.5.2 Test 3 202
4.4.5.3 Test 13 203
4.4.6 Tests 18, 19, 23, and 21 206
4.4.6.1 Tests 18 and 19 206
4.4.6.2 Tests 23 and 24 207

5. Assessing Test Difficulty: Data Analysis of Tests and Feedback Questionnaires 211
5.1 Setting the Scene for this Part of the Study 211
5.2 Reliability and Validity in Language Testing 213
5.3 Test and Questionnaire Administration and Data Analysis 216
5.4 Test Battery 1 (from Test 01 to Test 10) 220
5.5 Test Battery 2 (from Test 18 to Test 21) 240
5.6 Test Battery 3 (Tests 11, 12, 22, and 23) 251
5.7 Test Battery 4 (from Test 12 to Test 16) 260
5.7.1 From Test 13 to Test 16 262
5.7.2 Test 12 and the Perception of the Difficulty of the Tests in Battery 4 271
5.8 Test Battery 5 (Tests 17, 24 and 25) 278
5.9 Relationship between Test Takers' Performance, their Perception of Text/Task Difficulty and Results of the Quantitative Investigation 289

6. Conclusions 295

References 309
Appendix 321
Index 351