Grace Y. W. Tse

A Corpus-based Study of Proper Names in Present-day English

Aspects of Gradience and Article Usage
TABLE OF CONTENTS

Preface ix

1 Introduction 1
  1.1 Objectives 2
  1.2 Previous Approaches to the Study of Proper Names 2
    1.2.1 Philosophical Approach 3
    1.2.2 Linguistic Approach 3
    1.2.3 Onomastic Approach 9
  1.3 Methodology 11
  1.4 Description of Data 14
  1.5 Structure of the Book 15

2 A Semantic Study of English Proper Names 17
  2.1 Introduction 17
  2.2 A Classical Approach to the Study of English Proper Names 18
  2.3 Prototype Effects in English Proper Names 24
    2.3.1 Personal Names 26
    2.3.2 Place Names 27
    2.3.3 Organisation Names 30
    2.3.4 Names for Nationalities and Other Groups of People 31
    2.3.5 Names of Time-spans 32
    2.3.6 Names of Human Artefacts 33
    2.3.7 Technical Names 36
    2.3.8 Language Names 36
    2.3.9 Graded Membership in the Overall Category of English
    Proper Names 37
  2.4 Mixed Sub-categories of English Proper Names 38
  2.5 Conclusion 41

3 A Grammatical Study of Personal Names, Place Names and
   Organisation Names 43
  3.1 Introduction 43
  3.2 The Grammar of English Personal Names 45
3.2.1 Official Names of Persons
  3.2.1.1 Official Names Preceded by the Title (Without the Definite Article) 46
  3.2.1.2 Official Names Preceded by the Title and the Definite Article 50
  3.2.1.3 Official Names Followed by the Definite Article and the Common Noun 52
  3.2.1.4 Official Names Preceded by Honorific Adjectives 53
  3.2.2 Pure Title(s) as Unique Reference 55
  3.2.3 Nicknames and Epithets 56
  3.2.4 Some Grammatical Features of Minor Kinds of Personal Names 58

3.3 The Grammar of English Place Names
  3.3.1 Single-word Place Names 63
  3.3.2 Multi-word Place Names 68
    3.3.2.1 The Constituents of Multi-word Place Names 70
    3.3.2.2 Premodification in Multi-word Place Names 71
    3.3.2.3 Postmodification in Multi-word Place Names 79
    3.3.2.4 Other Structural Kinds of Multi-word Place Names 82

3.4 The Grammar of English Organisation Names
  3.4.1 Single-word Organisation Names 86
  3.4.2 Multi-word Organisation Names 88
    3.4.2.1 Premodification and Postmodification in Multi-word Organisation Names 91
    3.4.2.2 Other Structural Kinds of Multi-word Organisation Names 93

3.5 Conclusion 95

4 The Grammatical Factors Influencing the Choice between the Use and Omission of the Definite Article Preceding Multi-word Organisation Names: An Experimental Study 97
  4.1 Introduction 97
  4.2 Technique of Analysis: The Binary Logistic Regression Model 99
  4.3 Data Selection for the Experimental Study 101
  4.4 Preparing the Data for SPSS 102
  4.5 Inputting the Results of the Grammatical Classification to the SPSS Package 106
# Table of Contents

4.6 Statistical Analysis
  4.6.1 Model Selection 108
  4.6.2 Goodness of Fit 115
  4.6.3 Multicollinearity 121
  4.6.4 Analysis of Residuals 127
4.7 The Contribution of Each Explanatory Variable to the Model 137
4.8 Implication of Results for Statistical Analysis of Further Data 140
4.9 Conclusion 142

5 Validating the Logistic Model of Article Usage Preceding Multi-word Organisation Names 145
  5.1 Introduction 145
  5.2 Model Validation
    5.2.1 Testing the Validity of the Baseline Model on the Guardian Data Set 148
    5.2.2 Testing the Validity of the Baseline Model on the Daily Telegraph Data Set 150
    5.2.3 An 'Eyeball' Comparison of the Models across the Three Data Sets 152
  5.3 The Steps in Building an Overarching Logistic Model 153
    5.3.1 Variable Specification 154
    5.3.2 Selection of Essential Variables 158
    5.3.3 Interaction Assessment 159
    5.3.4 Building the Overarching Logistic Model 164
  5.4 Assessing the Performance of the Overarching Logistic Model 165
    5.4.1 Goodness of Fit of the Overarching Logistic Model 166
    5.4.2 Numerical Problems 166
    5.4.3 Analysis of Residuals 169
    5.4.4 Interpretation of the Regression Coefficients in the Overarching Logistic Model 178
  5.5 Conclusion 181

6 Predictive Analysis 183
  6.1 Introduction 183
  6.2 Gradience in Grammatical Factors 183
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.1</td>
<td>Factors that Favour the Use of the Definite Article</td>
<td>186</td>
</tr>
<tr>
<td>6.2.1.1</td>
<td>The PP Main Effect</td>
<td>186</td>
</tr>
<tr>
<td>6.2.1.2</td>
<td>The NP Main Effect</td>
<td>188</td>
</tr>
<tr>
<td>6.2.1.3</td>
<td>The NPH Main Effect</td>
<td>189</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Factors that Favour the Omission of the Definite Article</td>
<td>190</td>
</tr>
<tr>
<td>6.2.2.1</td>
<td>The PN Main Effect</td>
<td>191</td>
</tr>
<tr>
<td>6.2.2.2</td>
<td>The AC Main Effect</td>
<td>191</td>
</tr>
<tr>
<td>6.2.2.3</td>
<td>The NA Main Effect</td>
<td>192</td>
</tr>
<tr>
<td>6.3</td>
<td>Predictive Test</td>
<td>193</td>
</tr>
<tr>
<td>6.4</td>
<td>Conclusion</td>
<td>198</td>
</tr>
<tr>
<td>7</td>
<td>Conclusion</td>
<td>199</td>
</tr>
<tr>
<td>7.1</td>
<td>Summary and Theoretical Implications</td>
<td>199</td>
</tr>
<tr>
<td>7.1.1</td>
<td>A Cognitive Account of English Proper Names</td>
<td>199</td>
</tr>
<tr>
<td>7.1.2</td>
<td>The Internal Grammar of Personal Names, Place Names and Organisation Names: A Descriptive Linguistic Account</td>
<td>200</td>
</tr>
<tr>
<td>7.1.3</td>
<td>Gradience in Grammatical Factors in Determining Article Usage Preceding Multi-word Organisation Names: A Loglinear Analysis</td>
<td>201</td>
</tr>
<tr>
<td>7.2</td>
<td>Implications for Further Research</td>
<td>202</td>
</tr>
<tr>
<td>List of References</td>
<td></td>
<td>205</td>
</tr>
<tr>
<td>Appendix A</td>
<td>Details of the News Texts Selected from the British National Corpus</td>
<td>215</td>
</tr>
<tr>
<td>Appendix B</td>
<td>An Overview of Divided Article Usage Preceding English Place Names Discussed in the Literature</td>
<td>217</td>
</tr>
<tr>
<td>Appendix C</td>
<td>A List of Tags Used for Annotating the Corpus Data</td>
<td>221</td>
</tr>
<tr>
<td>Appendix D</td>
<td>Details of the News Texts Selected from the British National Corpus for the Predictive Test</td>
<td>227</td>
</tr>
<tr>
<td>Appendix E</td>
<td>The Data File for Predictive Test</td>
<td>229</td>
</tr>
<tr>
<td>Appendix F</td>
<td>Predictive Test: Summary Statistics</td>
<td>243</td>
</tr>
<tr>
<td>Index</td>
<td></td>
<td>247</td>
</tr>
</tbody>
</table>