

Overcoming Challenges in Corpus Construction

The Spoken British National
Corpus 2014

Robbie Love

Contents

<i>List of Tables</i>	ix
<i>List of Figures</i>	xi
<i>Foreword</i>	xiii
<i>Preface</i>	xvii
<i>Acknowledgments</i>	xviii
1 Introduction	1
PART I	
Before Corpus Construction: Theory and Design	7
2 Why a New Spoken BNC and Why Now?	9
3 Theoretical Challenges in Corpus Design	30
PART II	
During Corpus Construction: Theory Meets Practice	49
4 Challenges in Data Collection	51
5 Challenges in Transcription, Part I – Conventions	103
6 Challenges in Transcription, Part II – Who Said What?	136
7 Challenges in Corpus Processing and Dissemination	165

PART III

After Corpus Construction: Evaluating the Corpus 179

8 Evaluating the Spoken BNC2014 181

9 Conclusions and Further Construction Work 192

Index 201