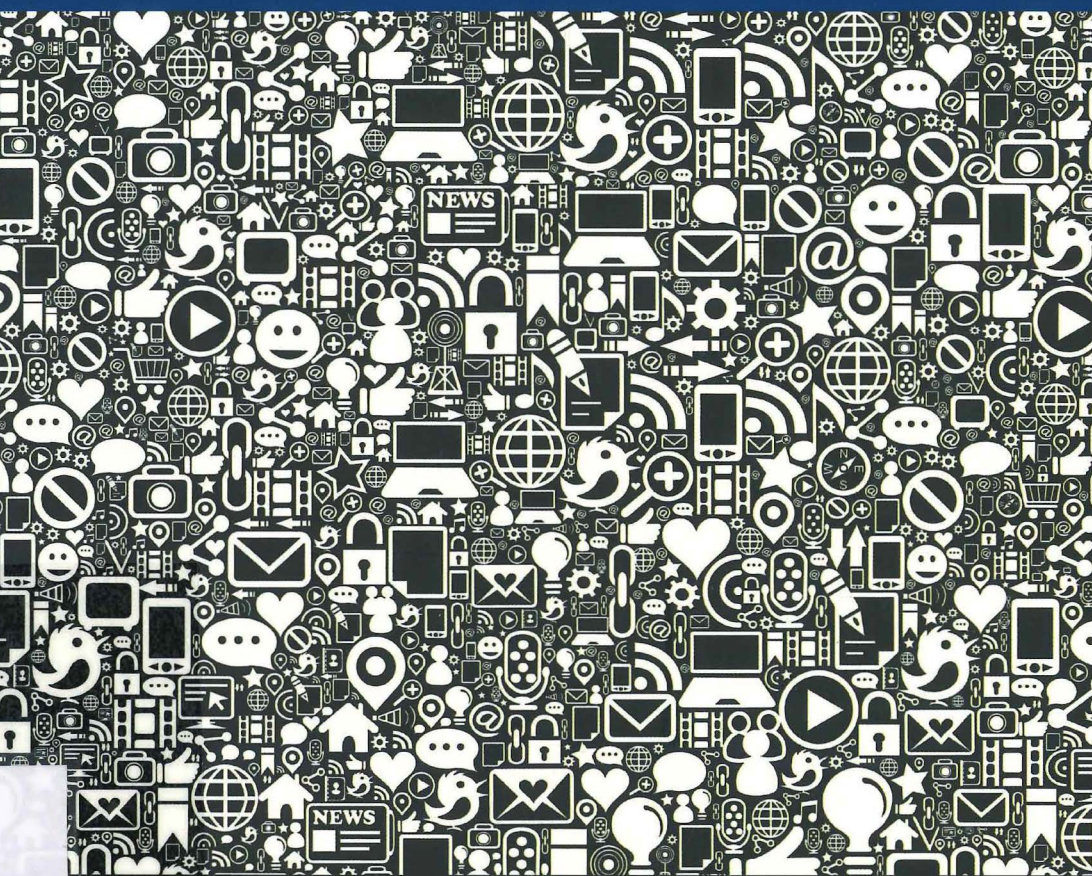


# RESEARCHING LANGUAGE AND SOCIAL MEDIA

A Student Guide

SECOND EDITION



RUTH PAGE, DAVID BARTON, CARMEN LEE,  
JOHANN WOLFGANG UNGER,  
AND MICHELE ZAPPAVIGNA

b1662029x  
012552823  
122607080

# Researching Language and Social Media

A Student Guide

*Second edition*



Ruth Page, David Barton, Carmen Lee,  
Johann Wolfgang Unger,  
and Michele Zappavigna

---

# Contents

---

<i>List of figures</i>	vi
<i>List of tables</i>	vii
<i>Acknowledgements</i>	viii
Introduction	1
1 What is social media?	5
2 What might a linguist say about social media?	25
3 What does it mean to research?	44
4 What are Internet research ethics?	57
5 Analysing discourse: Qualitative approaches	79
6 What are ethnographic approaches?	108
7 Carrying out a study of language practices on social media	131
8 Collecting social media materials for quantitative projects	145
9 Working with social media data: Quantitative perspectives	175
<i>Index</i>	201

---

## Figures

---

5.1 Factors that influence how much data to collect	84
5.2 Harriet the cat as a meme, photo by J W Unger	97
8.1 Continuum of discourse data. Adapted from Bednarek (2009, p. 19)	153
9.1 Screenshot of a concordance of the search term sorry using AntConc (Anthony 2005)	179
9.2 Word cloud generated from the content of Chapter 9 of this book	190
9.3 Visualisation of the Irish language blogosphere created by John Caulfield using Gephi software	193
9.4 An area graph for rebroadcast tweets following the death of Margaret Thatcher	194
9.5 Stacked area chart showing retweeted messages with and without a hashtag	195
9.6 A Twitter StreamGraph generated with the search word <i>linguistics</i>	196
9.7 Dialect map showing the distribution of <i>soda</i> , <i>coke</i> , and <i>pop</i> (Russ 2012)	197

---

# Tables

---

---

1.1	Timeline of selected social media platforms	7
1.2	Summary of Kaplan and Haenlein's (2010) factors for classifying social media types	12
1.3	Herring's (2007) medium and situation factors	12
2.1	Adapted from Herring's (2013) "Four levels of CMDA" and "Multimodal communication as an additional level of CMDA"	39
2.2	A selected summary of linguistic studies that examine social media sites	40
3.1	Methods for gathering, eliciting, analysing, and presenting data	55
8.1	Examples of linguistic and social variables	150
8.2	Examples of independent and dependent variables	150
8.3	Examples of different sorts of social media corpora	160
8.4	Examples of metadata about a post and a user account	165
9.1	An example of raw frequency in relation to word length	177
9.2	Top ten word frequencies in Birmingham blog corpus and HERMES Twitter corpus	184