

OXFORD

# Borrowed Words

*A History of  
Loanwords  
in English*



PHILIP DURKIN

# **Borrowed Words**

*Also by Philip Durkin*  
The Oxford Guide to Etymology

# Borrowed Words

A History of Loanwords  
in English

Philip Durkin

OXFORD  
UNIVERSITY PRESS

**OXFORD**

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,  
United Kingdom

Oxford University Press is a department of the University of Oxford.  
It furthers the University's objective of excellence in research, scholarship,  
and education by publishing worldwide. Oxford is a registered trade mark of  
Oxford University Press in the UK and in certain other countries

© Philip Durkin 2014

The moral rights of the author have been asserted

First Edition published in 2014

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in  
a retrieval system, or transmitted, in any form or by any means, without the  
prior permission in writing of Oxford University Press, or as expressly permitted  
by law, by licence or under terms agreed with the appropriate reprographics  
rights organization. Enquiries concerning reproduction outside the scope of the  
above should be sent to the Rights Department, Oxford University Press, at the  
address above

You must not circulate this work in any other form  
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press  
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data  
Data available

Library of Congress Control Number: 2013940283

ISBN 978-0-19-957499-5

Printed in Great Britain by  
CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and  
for information only. Oxford disclaims any responsibility for the materials  
contained in any third party website referenced in this work.

*For Kathryn*

# Acknowledgements

This book has arisen primarily from my work on *The Oxford English Dictionary*, and draws heavily on the collaborative research of my many very talented colleagues on the dictionary's staff. If the *OED* did not exist, this book, like very many others, would have been entirely impossible; if I had not had the good fortune to work in such a friendly and supportive environment, writing it would have been a much less pleasant experience.

I am greatly indebted to friends and colleagues from many corners of the world who have commented on drafts of either the whole or parts of this book, especially: Rhona Alcorn, Mark Chambers, Richard Dance, Anthony Esposito, Alan Kirkness, Roger Lass, Ursula Lenker, Serge Lusinian, Seth Mehl, Inge Milfull, Sara Pons-Sanz, Herbert Schendl, John Simpson, Janne Skaffari, Katrin Thier, and Edmund Weiner. Responsibility for any errors of course remains entirely my own.

I have also benefitted hugely from comments on papers containing research for this book from audiences at the Sixteenth International Conference on English Historical Linguistics (ICEHL), Pécs, Hungary; the Fifth International Conference on Historical Lexicography and Lexicology (ICHLL), Oxford; the Seventh International Conference on Middle English (ICOME), Lviv; the LIPP Symposium on linguistic change, Ludwig-Maximilians-Universität, Munich; HELLEX 3 (New Approaches in English Historical Lexis), Helsinki; the Colloquium 'Present and future research in Anglo-Norman', organized by the Anglo-Norman Dictionary team in Aberystwyth; and seminar audiences in Cambridge, Oxford, London (Westminster), Bamberg, Poznań, and Warsaw. I am very grateful to my hosts at all of these events for providing such valuable fora for working through and debating some of the ideas in the book.

That this book looks so good in its finished form owes a great deal to many talented colleagues at OUP, including Julia Steer, Jen Moore, Vicki Hart, and Briony Ryles. I consider myself, and this book, particularly fortunate to have benefitted from the generous and expert advice of John Davey, who retired from OUP at Easter 2013; his editorial skills are legendary, as is his contribution to the world of linguistics, and it is a great privilege to have worked with him on this project and others.

# Concise Contents

<i>Full Contents</i>	ix
<i>List of Figures</i>	xvii
<i>List of Abbreviations</i>	xx

## **PART I INTRODUCTION**

1 Introducing concepts	3
2 Introducing the data	22

## **PART II EARLY CONTACTS IN CONTINENTAL EUROPE AND BRITAIN**

3 Historical and cultural background to c.1150	53
4 Very early borrowings into Germanic	66
5 Old English in contact with Celtic	76

## **PART III OLD ENGLISH AND PROTO-OLD ENGLISH IN CONTACT WITH LATIN**

Introduction to part III	97
6 An overview of Latin loanwords in Old English	99
7 Interrogating the data from chapter 6	120
8 Methodologies: sound change; word geography; loanwords versus semantic borrowing	143
Conclusions to part III	168

## **PART IV SCANDINAVIAN INFLUENCE**

Introduction to part IV	171
9 Introduction to Scandinavian loanwords in English	173
10 Identifying Scandinavian borrowings and assessing their impact	190
Conclusions to part IV	220

**PART V BORROWING FROM FRENCH AND LATIN  
IN MIDDLE ENGLISH**

Introduction to part V	223
11 Exploring the contact situation and identifying loans	227
12 Quantifying French and Latin contributions to Middle English	254
13 Example passages from English and multilingual texts	281
Conclusions to part V	296

**PART VI LOANWORDS INTO ENGLISH AFTER 1500;  
HOW BORROWING HAS AFFECTED THE LEXICON**

Introduction to part VI	299
14 Borrowing from Latin and French after 1500	305
15 Loanwords from other languages: test cases	350
16 Long-term effects of loanwords on the shape of the English Lexicon	400
17 General conclusions and pointers for further investigation	424
<i>References</i>	429
<i>General Index</i>	455
<i>Word Index</i>	467

The companion website for the book is available at  
[www.oup.co.uk/companion/durkin2](http://www.oup.co.uk/companion/durkin2)

# Full Contents

<i>List of Figures</i>	xvii
<i>List of Abbreviations</i>	xx

## PART I INTRODUCTION

<b>1 Introducing concepts</b>	<b>3</b>
1.1 A first illustration of the part played by loanwords in the vocabulary of modern English	4
1.2 Some initial definitions of terms	7
1.2.1 Periods in the history of English	7
1.2.2 Types of lexical borrowing; borrowing and code switching; borrowing and imposition	8
1.3 Some different approaches to studying lexical borrowing	11
1.4 On evidence and hypotheses	13
1.5 What constitutes the vocabulary of English?	15
<b>2 Introducing the data</b>	<b>22</b>
2.1 Assessing input from different languages in the vocabulary of modern English	22
2.2 Examining loanwords in the high-frequency vocabulary of modern English	34
2.3 Assessing the impact of borrowing on the 'basic' vocabulary of English	41
2.4 Some implications of this data for the shape of this book	45

## PART II EARLY CONTACTS IN CONTINENTAL EUROPE AND BRITAIN

<b>3 Historical and cultural background to c.1150</b>	<b>53</b>
3.1 The Germani at the dawn of their recorded history	53
3.2 The Germani on the continent in later Roman times	54
3.3 Britain before the Romans	55

3.4 Roman Britain and its linguistic situation	56
3.5 From the Anglo-Saxon 'Settlement' to the first Christian centuries	59
3.6 The influence of Latin after the conversion	62
3.7 Anglo-Saxons and Scandinavians	64
3.8 The Norman Conquest	65
<b>4 Very early borrowings into Germanic</b>	<b>66</b>
4.1 Language families and comparative reconstruction	66
4.2 Very early borrowings from Celtic	69
4.3 Very early borrowings from Latin	72
<b>5 Old English in contact with Celtic</b>	<b>76</b>
5.1 Lexical borrowings from Celtic into Old English	77
5.2 The evidence of personal names and place names	81
5.3 A comparison: borrowing from Celtic into French	83
5.4 The hypothesis of structural borrowing from Celtic in English ('the Celtic hypothesis')	87
5.5 Epilogue: later lexical borrowing from Celtic languages	91
5.6 Conclusions	94
 <b>PART III OLD ENGLISH AND PROTO-OLD ENGLISH IN CONTACT WITH LATIN</b>	
Introduction to part III	97
<b>6 An overview of Latin loanwords in Old English</b>	<b>99</b>
6.1 Estimating the scale of the contribution	99
6.2 Earlier and later borrowings	101
6.2.1 Identifying earlier and later borrowings	101
6.2.2 Characteristics of earlier and later borrowings	103
6.2.3 Attempts to distinguish chronological and geographical layers of borrowing among the early loanwords	104
6.3 The loanwords	105
6.3.1 Early borrowings (to c. AD 650)	107
6.3.2 Some cases where an early date has often been suggested but is less certain	114
6.3.3 Some later loanwords (probably after AD 650)	116

<b>7 Interrogating the data from chapter 6</b>	<b>120</b>
7.1 Concerns about etymologies	120
7.2 Derivatives and compounds of loanwords	121
7.3 Uncertain cases of derivation or independent borrowing	123
7.4 Problems concerning learned borrowings	123
7.5 Assessing the influence of Latin loanwords in Old English	129
7.5.1 Survival in modern English	129
7.5.1.1 Survivals from probable earlier borrowings	131
7.5.1.2 Survivals from later borrowings	131
7.5.2 Word frequencies and textual distribution	132
7.5.3 Loanwords showing basic meanings	137
7.5.4 A closer examination of borrowed verbs	139
<b>8 Methodologies: sound change; word geography; loanwords versus semantic borrowing</b>	<b>143</b>
8.1 Evidence based on English and Latin sound change	143
8.1.1 A selective list of changes	144
8.1.2 Examples of Latin borrowings involving these changes	146
8.1.3 Examples involving changes in Latin as well as English	149
8.1.4 Changes of quantity reflected by late borrowings	154
8.1.5 How analogy can produce misleading forms: <i>proud</i> and <i>pride</i>	154
8.2 The problem of parallels/cognates in other Germanic languages	155
8.2.1 Greek words, the hypothesis of the 'Danube mission', and some word histories that have shown frequent reappraisals	158
8.3 Semantic borrowing	161
8.3.1 Cases where existing words acquire a new meaning	162
8.3.2 Cases where a new word is created	164
8.4 Word-geography, borrowing, and loan rendition reflected by the names of the days of the week in Old English and other Germanic languages	166
Conclusions to part III	168

**PART IV SCANDINAVIAN INFLUENCE**

Introduction to part IV	171
<b>9 Introduction to Scandinavian loanwords in English</b>	<b>173</b>
9.1 Areas of Scandinavian settlement in England	173
9.2 What is meant by 'early Scandinavian'	175
9.3 An illustrative example of some of the main issues: <i>they</i>	175
9.4 Periodization	179
9.4.1 Scandinavian borrowings attested in Old English	179
9.4.2 Borrowing as reflected in Middle English and in later sources	182
9.4.3 The likely date of borrowing and its context	187
<b>10 Identifying Scandinavian borrowings and assessing   their impact</b>	<b>190</b>
10.1 Identifying Scandinavian borrowings	190
10.2 Words distinct in sound from their native cognates	191
10.2.1 Absence of Old English /sk/ > /ʃ/	193
10.2.2 Absence of Old English /k/ > /tʃ/	194
10.2.3 Absence of Old English /g/ > /j/ and /gg/ > /ddʒ/	195
10.2.4 Absence of Old English development of Germanic *ai > ā	196
10.2.5 Germanic *au > ēa in Old English (open ē in Middle English)	197
10.2.6 *jj > gg in early Scandinavian	197
10.2.7 *ui > y in early Scandinavian	198
10.2.8 Loss of final nasals in early Scandinavian	198
10.3 Borrowings with different derivational morphology from a cognate	198
10.4 Words with no native cognate where formal grounds indicate borrowing	199
10.4.1 Absence of Old English /sk/ > /ʃ/ , /k/ > /tʃ/ , /g/ > /j/	199
10.4.2 Absence of Old English development of Germanic *ai > ā	200
10.4.3 Presence of early Scandinavian <i>i</i> -mutation (or <i>r</i> -mutation) of Germanic *eu	200
10.4.4 Presence of early Scandinavian development of <i>æ</i> as <i>ā</i>	200

10.5	Borrowings that preserve aspects of Scandinavian morphology	201
10.6	Cases where borrowing is argued for on grounds other than word form	201
10.7	Apparent substitution of Scandinavian sounds in native words and vice versa	208
10.8	Words that probably show semantic influence from Scandinavian	209
10.9	Regional distribution of Scandinavian words	211
10.10	Impact of Scandinavian borrowing on English core vocabulary	213
10.11	The relationships of Scandinavian-derived lexis with other English words	215
10.12	Borrowed word-forming elements and influence on other aspects of the grammatical system	217
	Conclusions to part IV	220

## **PART V BORROWING FROM FRENCH AND LATIN IN MIDDLE ENGLISH**

	Introduction to part V	223
<b>11</b>	<b>Exploring the contact situation and identifying loans</b>	<b>227</b>
11.1	Sources of data	227
11.2	Multilingualism in later medieval Britain	229
11.3	Borrowing from French, Latin, and French and/or Latin	236
11.3.1	Background	236
11.3.2	Borrowings solely from French	240
11.3.3	Borrowings solely from Latin	244
11.3.4	Borrowings from French and/or Latin	245
11.3.5	Likely mechanisms of borrowing	250
11.4	Reinforcement or re-borrowing of words borrowed in Old English	251
<b>12</b>	<b>Quantifying French and Latin contributions to Middle English</b>	<b>254</b>
12.1	Chronological breakdown	254
12.1.1	Latin, French, and Latin and/or French borrowings in English lexicography and lexicology	254

12.1.2	Analysis of the data of the <i>MED</i> and <i>OED3</i>	255
12.1.2.1	Semantic borrowing	264
12.1.3	Studies based on particular texts or corpora	265
12.2	Anglo-French and continental French	269
12.2.1	Methodology	269
12.2.2	Distinctively Anglo-French loanwords as documented in <i>OED3</i>	272
12.2.3	A test case: words identified as loanwords from continental French in the parts of <i>OED3</i> that overlap with <i>AND2</i>	275
12.2.4	Evidence in English lexicography for unrecorded Anglo-French words, forms, or meanings	277
12.2.5	Some conclusions	280
<b>13</b>	<b>Example passages from English and multilingual texts</b>	<b>281</b>
13.1	Example passages from Middle English texts with commentary	281
13.1.1	Passage 1: from the Final Continuation of the <i>Peterborough Chronicle</i>	281
13.1.2	Passage 2: from the <i>Ormulum</i>	283
13.1.3	Passage 3: from the <i>Ancrene Wisse</i>	284
13.1.4	Passage 4: from John Trevisa's translation of Ranulf Higden's <i>Polychronicon</i>	286
13.1.5	Passage 5: from Caxton's Prologue to <i>The Boke of Eneydos</i>	288
13.2	Some examples from multilingual texts and texts not in English	290
	Conclusions to part V	296
 <b>PART VI LOANWORDS INTO ENGLISH AFTER 1500; HOW BORROWING HAS AFFECTED THE LEXICON</b>		
	Introduction to part VI	299
<b>14</b>	<b>Borrowing from Latin and French after 1500</b>	<b>305</b>
14.1	The development of written English after 1500	305
14.1.1	The expanding functions of English; changes in relationships between writing in English, Latin, and French	306
14.1.2	Stylistic developments in the written language	307

14.1.3	A re-examination of the data surveyed in chapter 2 in light of these factors	310
14.1.4	Attitudes towards loanwords in English	316
14.2	Morphological differentiation between Latin and French borrowings	320
14.3	Latinate spelling forms and the respelling or remodelling of earlier borrowings	325
14.4	Affixes of Latin and French origin in English word formation	327
14.5	Test cases: selected word families in English	332
14.6	Continued semantic borrowing shown by earlier loanwords	335
14.7	Increasing word frequency over time	336
14.8	Modern scientific formations from elements ultimately of Latin and Greek origin	340
14.8.1	Formations in modern vernacular languages from Latin or Greek elements	341
14.8.2	Scientific Latin	344
14.9	Neoclassical compounding in English and other modern vernaculars	346
14.10	Summary and conclusions	347
<b>15</b>	<b>Loanwords from other languages: test cases</b>	<b>350</b>
15.1	Introduction	350
15.2	Loanwords from other European languages	354
15.2.1	Dutch (and Afrikaans) and Low German	354
15.2.2	High German and Yiddish	360
15.2.3	Spanish and Portuguese	364
15.2.4	Italian	369
15.2.5	Distinguishing Romance inputs in Early Modern English	372
15.2.6	Greek	375
15.2.7	Russian	378
15.2.8	Recent loanwords from French revisited	379
15.3	Loanwords from languages from outside Europe	383
15.3.1	Arabic	383
15.3.2	Hebrew	385
15.3.3	Languages of South Asia	386
15.3.4	Malay	389

15.3.5 Chinese	391
15.3.6 Maori	392
15.3.7 Japanese	395
15.4 Some common themes	398
<b>16 Long-term effects of loanwords on the shape of the English Lexicon</b>	<b>400</b>
16.1 Some key questions	400
16.2 The available research tools	401
16.3 Some test cases	405
16.3.1 Test case 1: loanwords among the 100-meaning 'Leipzig-Jakarta List of Basic Vocabulary'	405
16.3.2 Test case 2: the senses	411
16.3.3 Test case 3: the physical world	415
16.4 Conclusions	419
<b>17 General conclusions and pointers for further investigation</b>	<b>424</b>
<i>References</i>	429
<i>General works cited by author</i>	429
<i>Dictionaries, reference works, and databases cited by title</i>	452
<i>General Index</i>	455
<i>Word Index</i>	467

# List of Figures

2.1	Totals of loanwords from the 25 most prolific inputs in <i>OED3</i> .	25
2.2	The most frequent donor languages, arranged in five charts graded by scale.	26–7
2.3	All items in <i>OED</i> classified by origin. (i) <i>OED</i> headword entries only. (ii) All lexical items in <i>OED</i> .	29
2.4	Loanwords from French, Latin, and French and/or Latin in parts of <i>OED3</i> so far completed, arranged chronologically.	33
2.5	New words of all origins as reflected by parts of <i>OED3</i> so far completed, arranged chronologically.	34
2.6	Loanwords from French, Latin, and French and/or Latin as a proportion of all new words, as reflected by parts of <i>OED3</i> so far completed.	35
2.7	(Indirect) origins of 1,000 most frequent words in the <i>BNC</i> .	38
2.8	Totals of loanwords from Scandinavian, French, Latin, and French and/or Latin in the 1,000 most frequent words in the <i>BNC</i> , arranged chronologically.	39
2.9	Loanwords from Scandinavian, French, Latin, and French and/or Latin in the full WOLD meaning list, arranged chronologically.	43
4.1	The major Germanic languages.	67
5.1	Loanwords from Irish, as reflected by <i>OED3</i> (A–ALZ and M–RZ).	94
10.1	Divergent developments from a shared base in early Scandinavian and in English, and convergence in Middle English as a result of borrowing.	207
12.1	Words from French, Latin, and French and/or Latin as a proportion of all headword entries in the <i>MED</i> .	255
12.2	Words from French, Latin, and French and/or Latin as a proportion of all headword entries first recorded in the Middle English period in the <i>MED</i> .	256
12.3	Words from French, Latin, and French and/or Latin as a proportion of all headword entries first recorded in the Middle English period in <i>OED3</i> .	257
12.4	Absolute numbers of new words from each source per half-century (in <i>OED3</i> , A–ALZ, M–R).	258
12.5	Absolute numbers of French and Latin loanwords per half-century, combined totals (in <i>OED3</i> , A–ALZ, M–R).	259

12.6	Totals of all <i>OED3</i> headwords (of any origin) with first dates in each half-century.	260
12.7	Borrowings from each source as a percentage of all new words per half-century.	261
12.8	Borrowings from French and Latin as a percentage of all new words per half-century, combined totals.	261
12.9	Numbers of new words from each source per half-century as reflected in the 1,000 most frequent words in the <i>BNC</i> .	263
VI.1	Loanwords as a proportion of all new words.	300
VI.2	Loanwords from Latin and French, and loanwords from all other sources, as a proportion of all new words.	301
14.1	Loanwords from French, Latin, and French and/or Latin in parts of <i>OED3</i> so far completed, arranged chronologically, 1300–present.	310
14.2	Loanwords from French, Latin, and French and/or Latin as a proportion of all new words, as reflected by parts of <i>OED3</i> so far completed, 1300–present.	311
14.3	Printed books published per year, based on <i>ESTC</i> data.	313
14.4	Totals of all new words in <i>OED3</i> in the alphabetical range A–ALZ.	315
14.5	Verbs formed from Latin past participial stems.	321
14.6	Cases where there is an earlier parallel ultimately from the same Latin verb.	323
14.7	Words ending in <i>-or</i> probably borrowed from a Latin word ending in <i>-or</i> .	324
14.8	Words ending in <i>-ory</i> and <i>-ary</i> borrowed directly from Latin.	325
14.9	The proportion of words of French and/or Latin origin among high-frequency spellings in Helsinki Corpus data for Middle English and Early Modern English, alongside a comparable sample from the <i>BNC</i> .	338
14.10	The proportion of words of French and/or Latin origin among high-frequency spellings in Helsinki Corpus data for 1420–1500 and 1500–70.	339
15.1	Loanwords from Dutch, Low German, and Afrikaans, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	355
15.2	Loanwords from German, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	362
15.3	Loanwords from Yiddish, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	364
15.4	Loanwords from Spanish, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	366
15.5	Loanwords from Portuguese, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	369
15.6	Loanwords from Italian, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	371

---

15.7	Loanwords from Greek, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	376
15.8	The proportion of Greek loanwords that have entered English partly via Latin, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	376
15.9	Loanwords from Russian, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	378
15.10	Loanwords from Arabic, as reflected by <i>OED3</i> (A–ALZ and M–RZZ), 1500–present.	384
15.11	Loanwords from Hebrew, as reflected by <i>OED3</i> (A–ALZ and M–RZZ), 1500–present.	386
15.12	Loanwords from languages of South Asia, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	387
15.13	Loanwords from Malay, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	390
15.14	Loanwords from Chinese, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	391
15.15	Loanwords from Maori, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	395
15.16	Loanwords from Japanese, as reflected by <i>OED3</i> (A–ALZ and M–RZZ).	398

# List of Abbreviations

(See References for full details.)

<i>AFW</i>	<i>Altfranzösisches Wörterbuch</i>
<i>ALD</i>	<i>The Advanced Learner's Dictionary of Current English</i> (2nd edn., 1963; see also <i>OALD</i> )
<i>AND</i>	<i>The Anglo-Norman Dictionary</i>
<i>BNC</i>	<i>British National Corpus</i>
<i>COCA</i>	<i>Corpus of Contemporary American English</i>
<i>DEAF</i>	<i>Dictionnaire étymologique de l'ancien français</i>
<i>DMF</i>	<i>Dictionnaire de moyen français</i>
<i>DMLBS</i>	<i>Dictionary of Medieval Latin from British Sources</i>
<i>DNZE</i>	<i>The Dictionary of New Zealand English</i>
<i>DOE</i>	<i>The Dictionary of Old English</i>
<i>DOST</i>	<i>A Dictionary of the Older Scottish Tongue</i>
<i>DSAE</i>	<i>A Dictionary of South African English on Historical Principles</i>
<i>ECCO</i>	<i>Eighteenth-Century Collections Online</i>
<i>EDD</i>	<i>The English Dialect Dictionary</i>
<i>EEBO</i>	<i>Early English Books Online</i>
<i>ESTC</i>	<i>English Short Title Catalogue</i>
<i>FEW</i>	<i>Französisches etymologisches Wörterbuch</i>
<i>GSL</i>	<i>A General Service List of English Words</i>
<i>HTOED</i>	<i>Historical Thesaurus of the Oxford English Dictionary</i>
<i>MED</i>	<i>Middle English Dictionary</i>
<i>OALD</i>	<i>The Oxford Advanced Learner's Dictionary</i> (see also <i>ALD</i> )
<i>ODEE</i>	<i>The Oxford Dictionary of English Etymology</i>
<i>OED</i>	<i>The Oxford English Dictionary</i>
<i>OLD</i>	<i>Oxford Latin Dictionary</i>
<i>RMLW</i>	<i>Revised Medieval Latin Word-List</i>
<i>SND</i>	<i>The Scottish National Dictionary</i>
<i>SOED</i>	<i>Shorter Oxford English Dictionary</i>
<i>TLF</i>	<i>Trésor de la langue française</i>
<i>TLL</i>	<i>Thesaurus Linguae Latinae</i>
<i>TOE</i>	<i>Thesaurus of Old English</i>

# Part I

## Introduction



# 1

## Introducing concepts

The topic of this book is how borrowed words have influenced the vocabulary of English over its history. A central theme is how the histories of individual words are intertwined firstly with linguistic history, that is to say with larger-scale trends and developments in the history of English; and secondly with external, non-linguistic history, that is to say with historical events and developments, such as the arrival of the Anglo-Saxons in Britain or the Norman Conquest.

In linguistics, the term ‘borrowing’ describes a process in which one language replicates a linguistic feature from another language, either wholly or partly. The metaphorical use of the word ‘borrowing’ to describe this process has some well-known flaws: nothing is taken away from what is termed the donor language, and there is no assumption that the ‘borrowing’ or ‘loan’ will ever be returned. In many ways, the idea of influence would be more appropriate. However, the term ‘borrowing’ has been firmly entrenched in linguistics as the usual term to describe this process since the nineteenth century, to the extent that most linguists no longer even think of it as a metaphor. This book does not attempt to change the fundamental terminology of the discipline, but keeps to the basic framework of borrowing, donors, recipients, and loans.

Words have both a form and a meaning. Either component can be borrowed. This book looks in particular at those cases where both the form and (at least some aspect of) the meaning of a word from another language have been borrowed into English. These are conventionally called loanwords. (For a more detailed account of terminology, see section 1.2.2.) The category of borrowed words is sometimes restricted to words of this type, but is often extended to include other categories, such as loan translations. For example,

in modern English we have the words *omnipotent* and *almighty* in roughly the same meanings. *Omnipotent* is a loanword from Latin *omnipotēnt-*, *omnipotēns*; both the word form and the meaning have been replicated in English. *Almighty* is probably a (very early) loan translation of the same Latin word, based on analysis of its component parts *omni-* ‘all’ and *potēns* ‘powerful, mighty’. This book deals primarily with loanwords, but other types such as loan translations will also be considered at various points in the historical narrative.

Chapter 1 introduces some key concepts in the study of loanwords, and will give an overall impression of the impact of loanwords on the vocabulary of modern English. Chapter 2 looks at the proportions of borrowings in the lexis as a whole (as reflected by the wordlist of the *Oxford English Dictionary*), among the high-frequency words of contemporary English, and among those words that realize the most basic meanings. Additionally, it introduces some of the contact situations in the history of English that have led to the heaviest and most significant episodes of lexical borrowing.

### **1.1 A first illustration of the part played by loanwords in the vocabulary of modern English**

Loanwords make up a huge proportion of the words in any large dictionary of English. They also figure largely in the language of everyday communication and some are found even among the most basic vocabulary of English. Exact figures and percentages are problematic for various reasons, which are examined in the course of this chapter and in chapter 2, and which are investigated from various different points of view throughout this book. However, one easy way of illustrating the pervasive nature of loanwords in the vocabulary of modern English is simply to look at some passages from different types of contemporary writing. The following five passages are taken from (i) a book written for young children, (ii) a piece of popular fiction, (iii) a playwright’s diary (as prepared for publication), (iv) a scholarly monograph by a historian, and (v) a scientific research paper. I have underlined all loanwords (except for proper names), and all words that have been formed within English from loanwords.

‘Charlie’, she says, ‘they look like fish fingers to me, and I would never eat a fish finger.’  
‘I know that, but these are not fish fingers. These are ocean nibbles from the supermarket under the sea—mermaids eat them all the time.’

(Lauren Child *I will not ever never eat a tomato* (2000).)

‘Have you ever been to York Minster?’ Archie broke the silence, not taking his eyes off the glossy front-door.

Kate looked at him, irritated that his thoughts could be elsewhere. ‘A while ago.’

‘And?’ He turned to her. ‘What did you think?’

She shrugged. ‘I don’t know. It was good. What has it got to do with us being here?’

‘How good? Was it very good? Fairly good? Do you think it was better than Winchester? Did you eat in the café?’

(Claire Peate *Headhunters* (2009) 264–5.)

Switch on the radio after supper and catch most of Elgar’s First Symphony, music which invariably transports me back to boyhood<sup>1</sup> and walking up Headingley Lane on a summer evening after a concert in Leeds Town Hall. The evocative power of music is, I suppose, greatest when heard in live performance. This is a recording but it still casts a spell because I have come on it by accident. Had I put on the recording myself the spell would have been nowhere near as powerful because self-induced. Why this should be I can’t think, though doubtless Proust would know.

(Alan Bennett *Untold Stories* (2006) 253.)

To see the Host, however fleetingly, was a privilege bringing blessing. Those robbed of this privilege by misfortunes such as poor eyesight might be rescued by heavenly intervention. Conversely, the sacrilegious might be deprived of the ability to see the Host which they profaned.

(Eamon Duffy *The Stripping of the Altars* (1992) 101.)

Unique-event mutations inferred from binary marker data were used to condition the possible trees but otherwise did not contribute to the likelihood. Population splitting was modeled under strict fission with no subsequent background migration. Population growth was modeled as an exponential from an initially constant effective population size.

(*Molecular Biology and Evolution* 19 (2002) 1011=Weale et al. 2002  
in main references section.)

These short extracts illustrate some key aspects of the effects of borrowing on the vocabulary of English over time. A high proportion of the words in use in everyday, non-technical conversation in modern English are not borrowed and have either been in English right back to its Germanic origins, or have been formed within English. However, it would be very difficult to conduct any sort of conversation in modern English without using some loanwords and, crucially, you would need a very good knowledge of English etymology to do so successfully. People with a good

<sup>1</sup> Some scholars consider that *boy* shows a loanword, perhaps from French, but this is one of many etymologies that are uncertain and disputed.

knowledge of French and/or Latin will probably be able to guess at a lot (but not all) of the words that have been borrowed from these languages, because their form and meaning remain relatively close to those in the donor languages; many other loanwords are much more difficult to spot, including the considerable number of borrowings from early Scandinavian found in everyday English (part IV will look in detail at the reasons for this).

As these extracts also illustrate, more formal language in modern English and/or more academic topics of discussion generally involve using a higher proportion of borrowed words than more casual everyday conversation. These are chiefly words borrowed from French and/or Latin, or words formed ultimately from elements that come from Latin or Greek. This is because more formal registers of modern English, and the specialist vocabularies of academia and technology, show many more borrowings from these sources. How this situation developed historically will be a major focus of parts V and VI. Scholars often speak about a marked stratification in the vocabulary of modern English: there is an everyday vocabulary, which itself shows many borrowings from other languages, and then there is a more learned and formal vocabulary, which shows huge influence from French and Latin in its basic constituents and in its derivational processes. This sort of binary division of the vocabulary of English is probably somewhat oversimplistic, but the concept at least provides a useful framework.

Another useful concept in this context is dissociation: as a result of borrowing, many semantic fields in English show formally unrelated words for related concepts. For example, the usual adjective corresponding to *mouth* shows the completely unrelated word form *oral*; compare the situation in modern German, in which the relationship remains clear between *Mund* ‘mouth’ and *mündlich* ‘oral’; the vocabulary of German is often said to be relatively consociated compared with that of English.<sup>2</sup>

<sup>2</sup> For this famous example, see Leisi and Mair (1999) 51; the term *dissociation* originated (as German *Dissoziation*) in the first edition of this work, Leisi (1955). In the study of English in Germany it has been common to observe, as Leisi did, that English shows a much more dissociated vocabulary than modern German. It is unclear how far empirical work supports this observation (compare Sanchez (2008), Sanchez-Stockhammer (2009)), but quite aside from this question of contrasting the vocabulary of English and German, the concept of dissociation is useful in highlighting how the effects of borrowing have radically transformed the correspondence (or lack of it) between word forms and meaning relations in many semantic fields in English.

## 1.2 Some initial definitions of terms

### 1.2.1 Periods in the history of English

The history of the English language is conventionally divided into four main periods, Old English, Middle English, Early Modern English, and Later Modern English. This periodization reflects some major changes in the grammar, pronunciation, and vocabulary of English which coincide very approximately with the transition from one period to the next. However, the changes in question were not abrupt and did not occur all at the same time. Hence we can only speak of very gradual transitions between major periods in the history of the language; any precise chronological division between periods is necessarily arbitrary. Additionally, different scholars take different views on where best to place the boundary between periods.

In this book I use the same periodization that is used in the new edition of the *Oxford English Dictionary (OED3)*: Old English denotes the period up to *c.* 1150 (just under a century after the Norman Conquest); Middle English from *c.* 1150 to *c.* 1500 (not long after the introduction of the printing press to Britain);<sup>3</sup> and modern English from *c.* 1500 onwards, now usually subdivided into Early Modern English, up to *c.* 1750 (or sometimes *c.* 1700), and Later Modern English for the period after this. Present-Day English is sometimes used to refer specifically to the language of recent decades.

The transition from Old English to Middle English is characterized by (among other things): a very gradual process of reduction in the use of distinct inflectional endings to convey grammatical relations and a greater reliance on word order; the gradual loss of grammatical gender; the beginning of a phase of borrowing from French and Latin with major consequences for the vocabulary of English; and (during the transitional period) some significant borrowing from Scandinavian languages.

Early Modern English is characterized by (among other things) a series of far-reaching changes in the pronunciation of vowels known collectively as the Great Vowel Shift and a gradual increase in the importance and influence (at first mostly in writing) of a standard form of language based largely on the English of London and the south-east of England.

<sup>3</sup> The terms 'early Middle English' (normally taken to be up to *c.* 1325) and 'late Middle English' (normally taken to be roughly *c.* 1400 to *c.* 1500, or sometimes *c.* 1375 to *c.* 1500) are often used to distinguish phenomena characteristic of the opening and concluding stages of the Middle English period.

The early stages of Later Modern English are marked by greater codification of the standard variety (as in grammars and dictionaries) and, linked with this, a strongly normative tradition, i.e. the increasing importance of comment on and rules for how English should and should not be used. Additionally, the Later Modern English period shows significant shifts in the ‘centres of gravity’ of English use worldwide, from a language still used chiefly in the British Isles, to one used as first language in a wide variety of territories around the world, and, increasingly, as first choice as a *lingua franca* in international communication worldwide.

### **1.2.2 Types of lexical borrowing; borrowing and code switching; borrowing and imposition**

Lexical borrowing occurs when the lexis of one language (commonly called the donor language or sometimes the source language) exercises an influence on the lexis of another language (commonly called the borrowing language or sometimes the receiving language), with the result that the borrowing language acquires a new word form or word meaning, or both, from the donor language.

This book concentrates on loanwords, which, as noted in the introduction to this chapter, result from the borrowing of a word form with its meaning (or a component of its meaning) from one language to another. Thus English *image* ‘artificial imitation or representation of something’ reflects a Middle English borrowing of French *image* in the same meaning. Similarly *friar* (Middle English *frere*) reflects a borrowing of French *frere* (modern French *frère*), but in this case only in a very particular meaning: in French this is the usual word in the meaning ‘brother’, but in English it denotes only a particular type of metaphorical ‘brother’ in a mendicant religious order.<sup>4</sup>

The other main types of lexical borrowing involve borrowing of meaning but not (directly) of word form, and can conveniently be referred to under the cover term *semantic borrowing*.<sup>5</sup> In some cases the structure of a word in the donor language is replicated by a new word in the borrowing language,

<sup>4</sup> For a detailed account of this example, see Durkin (2009) 3–7.

<sup>5</sup> The classic account of different types of lexical borrowing remains Haugen (1950), although the approaches and terminology adopted by different scholars vary considerably. For recent overviews, see Haspelmath (2009), Winford (2010).

e.g., as already noted, English *almighty* (Old English *almihtig*) was probably formed on the model of Latin *omnipotēns*; the components *all* and *mighty* of the English word match closely the components *omni-* ‘all’ and *potēns* ‘mighty, powerful’ of the Latin word. This is typically called a loan translation or calque. In other cases, an existing word in the borrowing language acquires a new meaning from a word in the donor language; e.g. Old English *brōwung* ‘suffering’ probably acquired the additional meaning ‘(Christ’s) passion’ by a process of analogy with the meanings of Latin *passiō* ‘suffering, (Christ’s) passion’. This is often termed a semantic loan.

Semantic borrowing is very characteristic of the Old English period, and these examples and others are discussed in section 8.3, where the terminology that is normally used to describe semantic borrowings by scholars of Old English is presented in more detail. This is an area where terminology differs considerably. It is also an area where it is notoriously difficult to be certain that borrowing has actually occurred. Semantic borrowing does not form part of the main focus of this book, except (i) for comparison with the borrowing of loanwords, and (ii) for the special case of loanwords that show continuing borrowing of new meanings from the original donor. For instance, *to present* (borrowed from French and/or Latin c.1300) acquires some further meanings from French much later, e.g. ‘to stage or put on (a play)’ in the sixteenth century (see section 14.6).

A further, intermediate, category of loan blends is sometimes distinguished, where a borrowed word is adapted or remodelled using material from the borrowing language. For instance, Old English *fēferfuge* ‘feverfew’ (the name of a plant traditionally used for medicinal purposes) shows a borrowing of Latin *febrifugia*, but with substitution of Old English *fēfer* ‘fever’ for Latin *febris* in the same meaning. Similarly *brownetta* ‘brunette’ (1582) shows a borrowing of Italian *brunetta* but with substitution of English *brown* for Italian *bruno*. Cases like this, where a word from a foreign language has clearly entered English, albeit with some remodelling, are included in the scope of this book.

Loanwords in English sometimes show remodelling or suffixation of the borrowed word stem with a suffix that signals the word class that the loanword belongs to in English. For instance, in Old English borrowed verbs all need to be accommodated morphologically to one of the Old English verb classes, hence Latin *plantāre* > Old English *plantian* ‘to plant’, showing the infinitive ending *-ian* of verbs of the second weak class (see further section 7.5.4 on this). In later English this is seldom obligatory,

although compare for instance adverbs of manner, which typically show suffixation in *-ly*. In Middle English and modern English, and especially in Early Modern English, there has been a tendency for the (borrowed) adjective-forming suffixes *-al* and *-ous* to be extended to some borrowed adjectives that do not show the equivalent ending in the donor language, hence *academical* (1549) < Latin *acadēmicus*, beside later *academic* (1579) from the same source, or *illustrious* (?1566) < Latin *illustris*. Sometimes words from other languages are directly incorporated into complex new words in English; this is particularly a feature of how words and word elements ultimately of Latin or Greek origin are drawn upon in the terminology of the modern sciences, as in *oleiferous* ‘yielding or bearing oil’ (1804) < Latin *oleum* ‘oil’ + the (ultimately borrowed) combining form *-iferous*. Chapter 14 looks in detail at some of the difficult questions of analysis and categorization posed by formations of this type.

Another important term encountered at various points in this book is code switching. This is a difficult and rather controversial topic. Most scholars consider that code switching occurs when bilingual or multilingual speakers mix elements from more than one language within a single act of communication, whether within a sentence or in successive sentences. Code switching and borrowing are distinct processes, although how distinct is a matter of some debate. One hotly contested issue is whether it is useful to call switches at the level of a single word code switches, or whether these should be regarded as borrowings, simply occurring in the language of bilingual or multilingual speakers; when single-word code switches and nonce borrowings are distinguished, criteria can vary.<sup>6</sup> Aside from the theoretical aspects of this question, it is important to note that when one bilingual or multilingual speaker is communicating with another, even if most of the communication is in one language, words (and larger units) from the other language(s) can readily be introduced. In some social and cultural circumstances this can be an important channel for words ultimately to enter the usage of monolingual speakers as well. This topic figures

<sup>6</sup> For fuller discussion and further references, see Durkin (2009) 173–7. On the question of how to treat singly occurring foreign-language words, important accounts from differing positions are offered by, on the one hand, Myers-Scotton (2002) or Thomason (2003), and, on the other, Poplack, Sankoff, and Miller (1988), Poplack and Meechan (1998), Poplack (2004), or (drawing on interesting empirical data) Poplack and Dion (2012); see also Gardner-Chloros (2010).

particularly prominently in part V of this book, when we examine English, French, and Latin in contact in the multilingual society of later medieval England.

A further theoretical distinction is sometimes drawn between borrowing and imposition, the latter term being used for the process by which speakers introduce new material into a language in the process of a shift from primary use of one language to primary use of another. This is typical in a situation of language death, where a community ceases to use one language in favour of another. This is considered in a little more detail in part IV, in the context of the Scandinavian contribution to the lexis of English.

Proper names, both place names and personal names, are excluded from the scope of this book. The methodology for studying borrowed names and name-forming elements is very different from that used in studying borrowed general lexis, and is best left to specialist studies.<sup>7</sup> However, in chapters 5, 9, and 10 the evidence of name studies is drawn upon to some extent for the light that it can shed on English loanwords from Celtic and Scandinavian languages.

### **1.3 Some different approaches to studying lexical borrowing**

Because loanwords are so pervasive in English, there are numerous different ways in which they can be approached and studied. We may look at how they have entered English, which languages they have been borrowed from, and at what times and in which contact situations. Alternatively, we may look at their impact on the structure of the lexis of English, examining how they enter into word-forming patterns within English, and/or what their impact is on the meaning relations between different words. Or it is possible to approach them primarily from the point of view of their stylistic or pragmatic effects in different text types from different periods. In this book the first, fundamentally etymological and historical, approach predominates, as is more or less traditional in most loanword studies. A good deal of attention is also given to the second approach, which, like the first approach, is based fundamentally on the study of individual word histories,

<sup>7</sup> On some of the main differences in the nature of the evidence see Durkin (2009) 266–83. For an introduction to the study of the place names of England see Cameron (1996). On the available resources for English personal names and surnames see McClure (forthcoming).

drawing generalizations from these to reveal something about the lexical history and development of English. The third, essentially stylistic, approach is rather different, being fundamentally based on the analysis of texts rather than of words across time, and requiring sensitivity to speakers' perceptions about words, as well as to the facts of their history (which are often entirely opaque to speakers, and are at best reflected indirectly through aspects of word structure or phonology). This approach figures much less prominently in this book, although some perspectives of this sort are examined at various points in the historical narrative, and especially in chapters 14 and 16.<sup>8</sup>

Another important issue in any historical survey is whether we are interested in investigating say Middle English borrowing from the point of view of Middle English, i.e. for what it tells us about what words speakers of Middle English used and how they used them, or for what it tells us about the historical background of the English used today, essentially a teleological approach. This book tries to keep an eye on both questions, but (unlike many surveys) attempts to maintain a distinction between the two; in particular, the surveys of loanwords in the high-frequency vocabulary of modern English and in the basic vocabulary of modern English in chapter 2 will be used as a point of reference throughout this book, in order to illuminate how borrowing in the past has shaped the everyday English of today.

A recent important trend in linguistic research has been to examine lexical borrowing in the context of broader issues of language contact and to attempt to classify the different sorts of linguistic borrowing that typify different sorts of contact situations.<sup>9</sup> In this book lexical borrowing, and specifically loanwords, are foregrounded throughout, and there is no attempt to offer an overview of all types of linguistic borrowing (syntactic, morphological, phonological, etc.) in the history of English.<sup>10</sup> However,

<sup>8</sup> See Fischer (2003) for a 'typology of typologies' of lexical borrowing. He distinguishes between approaches that are predominantly morphological, those that are predominantly semantic, and those that are predominantly sociolinguistic. This book concentrates chiefly on the morphological and semantic aspects, rather than the sociolinguistic.

<sup>9</sup> A very influential, albeit controversial, study of this type which takes a number of its case studies from episodes in the history of English is Thomason and Kaufman (1988). See also Thomason (2001) and the various contributions in Hickey (2010). The classic seminal work on language contact in general remains Weinreich (1953).

<sup>10</sup> For an important attempt at such a history up to the Renaissance, see Miller (2012).

other types of linguistic borrowing are touched on with regard to certain periods when it has been suggested that there was a significant discrepancy between lexical and non-lexical borrowing; this is especially (albeit controversially) the case with regard to early contact between English and Celtic, as will be considered in chapter 5.

Many recent studies of linguistic borrowing foreground sociolinguistic perspectives. In particular, a rather broad-brush categorization of the sociolinguistic relationship between any two languages in contact is often applied. If one language has a position of lower social and cultural prestige vis-à-vis another, it is said to be a substrate and in a substratal relationship with the other language. If it has higher prestige, it is a superstrate, in a superstratal relationship. And if the levels of prestige are roughly equal, the two languages are adstrates, in an adstratal relationship. In practice, the application of these terms varies considerably, with some scholars classifying as adstratal situations where the difference in prestige is relatively small, while others would classify the same situation as showing a relationship between substrate and superstrate. In this book, these terms will be used very sparingly and mostly in relation to early contact situations in which only the broadest indications of the nature of the sociolinguistic relationships between languages can be reconstructed. Fortunately, for most of the period covered by this book, we know quite a lot about the relevant contact situations and thus a fairly fine-grained analysis is possible. This is particularly true of loanwords in the Middle English and modern English periods. See further section 2.4 for an outline of the approach taken in this study.

#### **1.4 On evidence and hypotheses**

Lists of loanwords given in handbooks and histories of English can give the appearance of being simple statements of fact. It is important to realize that they are not: they are hypotheses, sometimes supported by evidence so secure that they are not in any real doubt, but very often based on much less secure foundations.

An entirely satisfactory, secure loanword etymology might show the following characteristics:

1. The supposed borrowing is first recorded later than the supposed donor (assuming that we have a dependable documentary record for each language in the relevant period).

2. The supposed borrowing shows form(s) entirely explicable from the form(s) of the supposed donor (allowing for later known processes in the borrowing language).
3. The supposed borrowing shows meaning(s) entirely explicable from the meaning(s) of the supposed donor as starting point.
4. There is a known historical context of language contact in which the borrowing could have occurred.
5. There is no alternative explanation for the supposed borrowing, or at least none that is as convincing as the assumption of borrowing from the supposed donor.

There can be problems with all of these criteria and not all of them are amenable to rigorous objectivity. As regards (1), our documentary record is often poor, or entirely absent, for both the borrowing language and the donor in the period when the borrowing is supposed to have occurred. As regards (2), not all form developments are entirely regular or easily explained. As regards (3), it is notoriously difficult to test whether a supposed semantic development is plausible.<sup>11</sup> As regards (4), our historical data is often sketchy, and we may also encounter considerable uncertainties about the likely historical setting if we do not know when or where a borrowing is likely to have occurred. As regards (5), there is often a good deal of subjectivity involved in deciding whether one etymology is more convincing than another, however hard we try to base our decisions firmly on the analysis of empirical data; also, scholars differ over how much priority to give to explanations that do not involve contact with other languages over ones that do, in other words, how much priority to give to endogeny over exogeny.

This book looks in some detail at problems of evidence and its interpretation that often make us much less certain than we would like about which words have been borrowed into English, from which languages, in which places, and at which times. This involves looking at some very complicated issues, but the reader who perseveres should have a much more informed understanding of some of the assumptions, hypotheses, and uncertainties that underlie the sometimes rather bland statements made about how many words English has borrowed from other languages. Through an emphasis

<sup>11</sup> For discussion and illustration of these last two points, see Durkin (2009) chapters 7 and 8.

on how we (think we) know what we do know, I hope also to illustrate how rich a field this is for further study.

### 1.5 What constitutes the vocabulary of English?

Another potential pitfall is to treat ‘the vocabulary (or lexis) of English’ as though it were an entirely unproblematic concept. Modern standard English, as used in Britain, the USA, and other majority English-speaking communities, is the product of a number of social, cultural, and technological factors operating over a long period of time. The rise of modern standard English as a written variety was closely bound together with the development of the printing industry in early modern England, as the choices made by printers from the available pool of variation in spelling forms and in vocabulary, especially core vocabulary, gradually converged on a particular set of norms. The spread of these linguistic choices beyond the printed medium was a slow process and owed a lot to normative trends in grammatical works and dictionaries in the eighteenth and nineteenth centuries. (Chapter 14 looks more closely at these processes.)

In the modern standard variety that resulted from these processes, lexical choices are often very constrained in the core vocabulary, especially among core grammatical items. Thus, *you are* is considered the only ‘correct’ form of the second-person singular of the verb *be*. *Thou art* may be familiar from the works of Shakespeare and his contemporaries even to those with no specialist knowledge of the history of English, but it is considered an archaism. Where it survives in regional speech, *thou art* is considered to be non-standard, and will therefore be avoided by most speakers coming from such areas when they are speaking in more formal environments; the same applies to other forms found in other varieties of English, such as *you be*, *you is*, *you am*, or *you’m*. This differentiation between standard and non-standard is the result of the application of social norms, which can themselves change. For instance, modern spoken British English shows a good deal of variation between standard *you were* and non-standard but very widespread *you was*; typically, *you was* is perceived as uneducated or ignorant, and avoided in formal contexts; however, in contemporary Britain it can be heard regularly in the media in spoken use by people who are widely respected as high achievers in various fields (sport, business, entertainment). In the future it may become more widely accepted even in formal contexts as social attitudes shift. To take another example, some varieties of English

(especially in the USA and Caribbean) have a pronoun of the type *you-all*, others do not; among those that do have it, in some varieties its use is restricted to addressing more than one person, but in others it can also be used to address just one person. A similar situation is found with *youse* (and a number of other forms) in some varieties of British and US English. Very many speakers who use *you-all* or *youse* when addressing other members of their own local communities will substitute *you* in more formal contexts, but the sociolinguistic situation is complex and fluid, and closely tied up with wider questions of regional identity and prestige.

At present, there is very little difference in core vocabulary between the major standard varieties of English worldwide. British and US English show some minor differences in grammatical vocabulary; e.g. in British English the past participle of *get* is invariably *got*, but in US English there is variation between *got* and *gotten*, depending on the semantic context. There is also some well-known variation in everyday non-grammatical vocabulary, such as *lift* versus *elevator*, *rubbish* versus *trash*. Similarly, in South African English *robot* is an everyday term for a traffic light. Today, smaller dictionaries are typically produced in editions specially tailored to different national varieties of English and reflecting these differences in lexical usage. However, this is a recent trend, linked to the growing prestige of different national varieties as markers of national identity, and the cases where usage is identical in each variety far outweigh the differences.

The differing incidence of loanwords of various origins is one of the more obvious ways in which varieties of English around the world do differ from one another. We will look at this topic more closely in chapter 15, but a first illustration can be gained by taking (very unscientifically) a passage from a contemporary South African short story and underlining all of the loanwords, in the same way as in section 1.1:

‘Don’t you want to see your grandson?’

‘Not if that moegoe has to be here.’

‘That moegoe is a part of our family now, whether we like it or not.’ The words slipped out unexpectedly.

‘Okay, I see how it is now.’

‘Ag, suit yourself then.’ Pauline returned to her daughter.

(Sean William O’Toole *The Marquis of Mooikloof and other stories* (2006) 19.)

Here, among numerous older borrowings that are common to most varieties of English, we have two distinctively South African English loanwords: *ag*,

an interjection borrowed from Afrikaans; and *moegoe*, a word meaning roughly 'bumpkin', borrowed into English probably immediately from Afrikaans and Isicamtho (a mixed language of urban South Africa), but of uncertain ultimate origin. Both words are included in the full *OED*, but labelled as restricted to South African English. Some words originally found in South African English have subsequently spread to other varieties, e.g. *commando* (< Portuguese), which is found in South African English from the late eighteenth century, denoting a body of irregular troops, but is rarely found in other varieties of English until the twentieth century, having become widely familiar in Britain as a result of the Boer War of 1899–1902, and then having acquired its modern meaning as a result of specific use by Winston Churchill in 1940 during the Second World War.

However, differences of vocabulary can also be found at the micro level, between different individuals. We all have slightly different individual vocabularies, both active, i.e. the words we use ourselves, and passive, i.e. the words that we understand but are unlikely to use ourselves. These differences in individual vocabulary reflect our different experiences, educational backgrounds, professional lives, interests, friendships, and other social and cultural factors. Words can then spread from the vocabulary of one person to that of another, as a result of direct social interaction or sometimes as a result of use in published writing or broadcasting, and so on. Intra-linguistic spread of this sort is strictly a type of lexical borrowing, although the term is more usually used specifically of borrowing between languages.

One way of approaching the topic of the differing vocabularies of different individuals is to consider the vocabulary of a particular specialist field. For instance, there is a large specialist vocabulary relating to wine. There are some terms that are widely familiar and are found even in dictionaries for learners like the *Oxford Advanced Learner's Dictionary (OALD)*, e.g. *corked*, describing wine which is spoiled because of decay in the cork; or allowing a wine to *breathe*, i.e. to begin to react with the air before it is drunk. They may be more familiar to some people than to others and some people may have a clearer idea of their meaning than others, but they are at least familiar to a large number of people. Others, like *laying down* fine wine to mature before it is drunk, or *drinking window*, denoting the period in which a mature wine is good to drink before it is past its peak, will be unfamiliar to many people, but well known to wine enthusiasts. Other people may be aware of these terms, but consciously avoid them, perceiving

them to be affected ‘snob’ wine language. None of these terms has entered the wine vocabulary of English as loanwords, but much of the specialist vocabulary of wine has. Some of these words have been in English for a long time and have even spread outside the specialist vocabulary of wine: e.g. *vintage* entered English (late in the Middle English period) as a specific term relating to wine, but developed various metaphorical uses in other fields and can nowadays be applied also to *vintage cars*, for instance. The word is fully naturalized in English and, although it was originally a borrowing from French, it differs considerably from the modern French word form, *vendange*. There are also very many French terms that are not at all widely known, but that are known to and used by many enthusiasts for French fine wine worldwide, e.g. *barrique*, *cépage*, *cuvée*, *élevage*, *en primeur*, *grand cru*, *grand vin*, *négociant*, *premier cru*, or *vigneron*. These all show little if any naturalization in pronunciation or morphology in English, and not all of them are found even in the largest English dictionaries, such as the *OED*. But they do all occur in specialist wine publications written in English and are definitely part of the active vocabulary of some English speakers and of the passive vocabulary of others. Some speakers will know some of these words but not others. Some will be equally familiar with the specialist terminology relating to German wine, e.g. *Anbaugebiet*, *Auslese*, *Bereich*, *Eiswein*, *Kabinett*, *Prädikat*, *Qualitätswein*, *Spätlese*, *Trockenbeerenauslese*, but others may not be. The minority group of English speakers familiar with such terminology is not homogeneous: it is widely spread geographically (wherever there are people who have a strong interest in French or German fine wine) and, although its members may tend to have some other interests in common, there will be others that they do not share, and they will belong to a range of different professions. Some may be fluent speakers of French and/or German, and their use of these words may show single-word code switches to French or German, but many will not be. In the approach taken in this book such words are regarded as loanwords, even if they are rare and non-naturalized ones.<sup>12</sup>

<sup>12</sup> In the tradition of most German lexicography and lexicology a distinction is made between *Lehnwörter*, ‘loanwords’, which are taken to show full phonological and morphological integration into German (the term is often restricted partly or wholly to very early loanwords, such as *Wein* or *Strasse*, on which compare part III) and *Fremdwörter*, ‘foreignisms’, the latter term being used for everything from loanwords which retain some aspect of their foreign-language morphology (e.g. in the way that they form the plural) to

Thus, when we speak about the vocabulary (or lexis) of a language, it can be useful to think of a (not very precisely defined) common core of basic vocabulary, including words in everyday use and function words. This generally shows relatively little variation within narrowly defined speech communities, or within standard varieties, such as modern-day standard English as used internationally. Beyond this we can attempt to identify numerous different (overlapping) vocabularies belonging to different specialisms (e.g. activities, professions), to different functions (e.g. academic discourse), or to different stylistic registers (formal, informal, etc.).<sup>13</sup> In each of these, some words will be frequent and known to most if not all speakers, others will be very rare and restricted to very small numbers of speakers. Borrowed words can belong to any of these categories: some, as already seen, are in everyday use in even the simplest discourse, while others are used only very occasionally in very restricted circumstances and only ever by certain speakers.

All of these factors need to be borne in mind when we consider questions about how much borrowing from various different languages the vocabulary of modern English shows. Are we looking at the wordlist of a small dictionary, concentrating on the core plus the most frequent words from specialist vocabularies? Or are we looking at the wordlist of a very large dictionary, which will include many words not in frequent use, many of which are only known by a small minority of speakers? We explore some of the practical implications of this in chapter 2.

If we introduce a diachronic dimension, things become more complicated again. As noted above, different varieties of English today can show variation even in the area of basic vocabulary and among grammatical function words, e.g. *you-all* and *youse*. Further back in the history of English such variation becomes much more typical. Many of our surviving Old English written documents rather mask dialectal variation, because they are written in a particular form of the West Saxon dialect that was in widespread use as

completely unassimilated items like those listed in this paragraph, and often embracing any loanwords from recent centuries, whatever their degree of integration. In practice, the two categories are impossible to distinguish consistently and reliably in all instances, and the distinction is not used in this book. See further Durkin (2009 139–40), and also Polenz (1967), Eisenberg (2011), and compare sections 7.4 and 14.8.1 of this book.

<sup>13</sup> An early form of this model was set out by Sir James Murray in his ‘General Explanations’ accompanying the first completed volume of the *OED* (Murray 1888).

a written variety in the late Old English period. However, our Middle English records are far more revealing about dialectal variation and enable us to reconstruct many different local systems of phonology, morphology, and syntax. In the Middle English period vocabulary varies hugely between different local varieties, even at the level of closed-class grammatical words: in chapter 9 we look at how the borrowed third-person plural pronoun *they* (and *their* and *them*) spread gradually between different dialects, by a process of internal borrowing from dialect to dialect. In a historical context like this, where there was so much variation even in core grammatical vocabulary, the concept of a common core of basic vocabulary can probably be meaningful for a particular dialect, but it is somewhat more difficult to speak of the core vocabulary of Middle English as a whole.

It is notoriously difficult to draw a distinction between two dialects of the same language and two different languages; the old observation that a language is a dialect with an army and a navy certainly has some merit, since a lot depends on questions of politics and national identity, connected with processes of linguistic standardization. A good example is presented by Scots in the sixteenth century. It had a distinct written form, with spelling conventions very different from those in the developing London-centred standard of England. It presented significant differences from the English of England (especially southern England) in pronunciation and in some areas of grammar, and, most obviously, it had very many vocabulary items not found in other varieties of English.<sup>14</sup> But, on the other hand, mutual comprehension was certainly possible, and many features were shared with neighbouring dialects of northern England. Over the following centuries as political union developed, so did the influence of the standard English of England on the English used in official functions and in more formal contexts in Scotland. Thus it is a very difficult question whether Scots constituted a separate language from the rest of English in the sixteenth century; if we decide that it was a different language, and that some rural dialects remain as the descendants of this, it is still more difficult to say at which subsequent point the language used in standard functions in Scotland became something that we would want to call English rather than Scots. Whether we decide to call English and Scots in the sixteenth century two different languages or two dialects of the same language, there are

<sup>14</sup> On the composition of the vocabulary of Older Scots, see especially Macafee and Anderson (1997), Macafee (1997), Macafee (2002).

interesting implications for how we talk about the vocabulary of English: either there was a period in which English had a neighbour with a great deal of vocabulary in common and easy mutual comprehensibility (especially in the border regions), or there were two different emerging standard varieties of English which showed very considerable differences even in aspects of their core vocabulary.

# 2

## Introducing the data

### 2.1 Assessing input from different languages in the vocabulary of modern English

Percentages are often quoted for the proportions of the vocabulary of modern English that are borrowed from French, Latin, Scandinavian languages, etc. As discussed in chapter 1, such percentages must be approached with extreme caution. Firstly, we have to bear in mind that such figures can only refer to a particular period: the proportions in contemporary English will not at all be the same as those in sixteenth- or seventeenth-century English, for example. Secondly, we must consider whose English we are talking about, as discussed in section 1.5. Thirdly, once we have decided which registers, varieties, etc. we want to take into account, we have the practical problem of arriving at a wordlist. Fourthly, once we have our wordlist, we have the problem of assessing and classifying etymologies, i.e. deciding which words are borrowed and which are not. This last problem is a major concern of this book.

Nonetheless, it can be useful to have some initial points of orientation, however much we may modify or qualify these later. The usual starting point for such discussions is the *OED*, comprising a wordlist of over 600,000 lexemes (including simplex words such as *butter*, compounds such as *butter knife*, and derivative formations such as *buttery* ‘rich in butter’) arranged under nearly 275,000 headword entries. Some of these words are now obsolete, but most are not. Although a total count of over 600,000 lexemes is vast, the *OED* should not be taken as a catalogue of every word or expression that ever is or has been used in English. Section 1.5 investigates how loanwords are one area (among many) in which the vocabulary of English can be hugely variable from one individual to another, and how a

common core of shared vocabulary shades off into hugely varied individual repertoires of lexical items depending on location, profession, interests, and individual experience. The *OED*'s selection criteria for including lexemes pay careful attention to questions of frequency of use and spread in different contexts of use, and its wordlist can be taken as a good reflection of those words that have had most impact on the lexicon of English since the mid-twelfth century (it generally includes Old English words only when they have survived into later use). No dictionary could ever list *all* English words, since the vocabulary is almost infinitely extendible by regular processes of word formation as well as by borrowing from other languages.

The *OED* has been compiled over a very long time: the first edition was published in fascicles (or instalments) between 1884 and 1928; supplements were added in 1933 and (in four volumes) between 1972 and 1986; and these were brought together with the main text in the Second Edition of 1989. A revised edition of the full text of the dictionary (*OED3*) has been in course of publication since 2000, and currently covers approximately one-third of the full text of the dictionary.

Fig. 2.1 presents the totals of loanwords from the most prolific donor languages as reflected by the two large contiguous alphabetical ranges of *OED3* that have so far been published, as of November 2012, comprising all of the letters M, N, O, P, Q, R, and A from A to ALZ. Approximately 92,500 dictionary entries fall into these alphabetical ranges, constituting the bulk of the *OED3* material so far published. For statistical purposes, I will not draw on shorter sequences of entries that have been published separately, such as *be*, *love*, or words beginning with *sub-* or *super-*, since these constitute a less representative cross-section of the English lexicon. However, for illustrative examples, I will draw freely on all of *OED3*, and also on unrevised parts of the dictionary.

The benefit of drawing statistical data from *OED3* alone is that the material has been edited over a relatively short period of time, using much the same set of resources for each word (e.g. dictionaries and databases for information about foreign languages, etc.), applying the same set of editorial principles and guidelines, and, just as importantly, employing the same style, so that data can easily be extracted computationally.<sup>1</sup> This portion of

<sup>1</sup> In compiling the data from *OED3* given here, as likewise in a number of other places in this book, I have made use of the sophisticated electronic tagging and search software of the computer system used by editors preparing the new edition of the *OED*. This is not

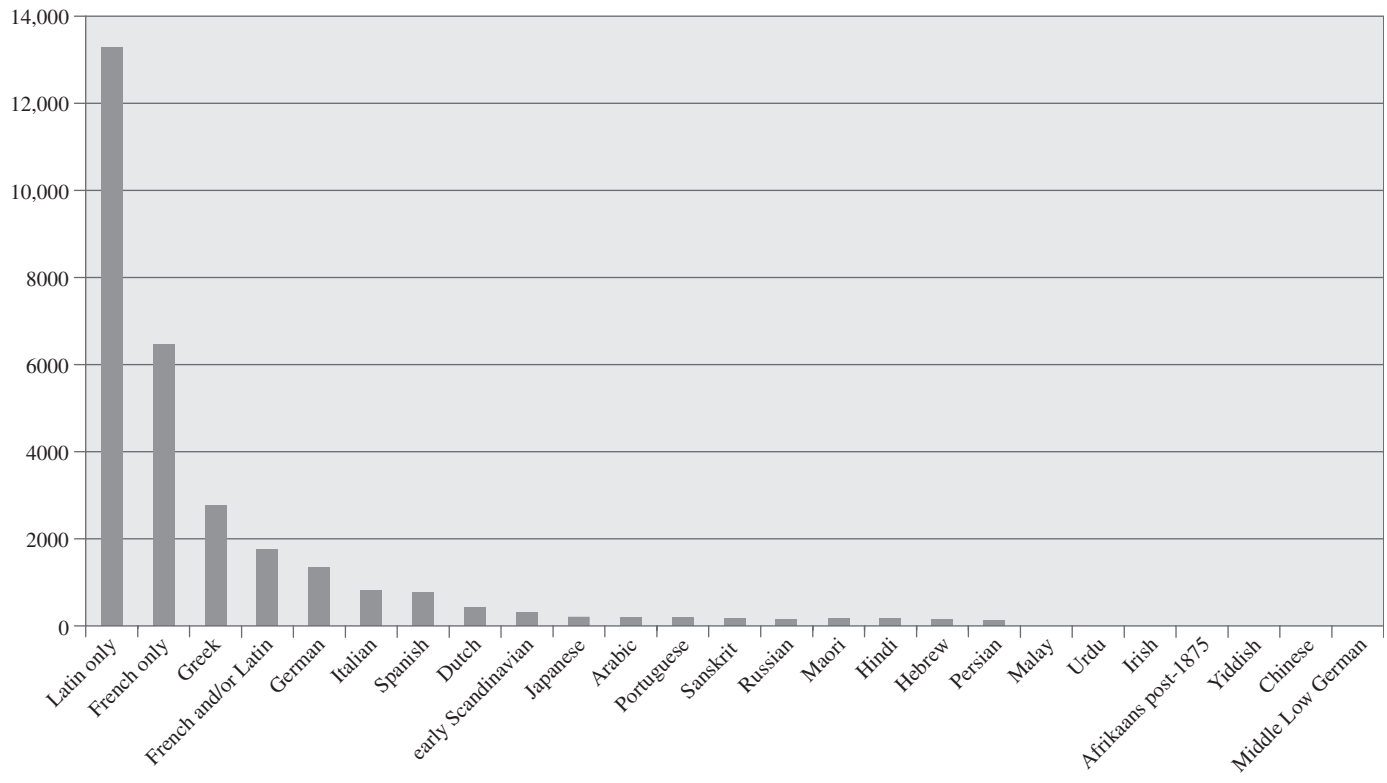
the dictionary also benefits from much more thorough treatment of recent change in the lexicon of English, including recent loanwords.<sup>2</sup> Additionally, any reader who wishes to compare the data presented here with data based on the whole of the *OED*, both revised and unrevised portions, can easily do so by consulting the timelines for loanwords from each donor language available at <<http://www.oed.com>>. For many of the languages discussed in this book, the lexicographical treatment of loanwords in *OED2* and *OED3* is very similar, and timelines based on the whole of *OED* differ very little in outline from those based only on *OED3* data. However, for some of the major donor languages, there are significant differences in treatment, reflecting issues that are discussed in detail in this book (especially in parts V and VI), and these do make a difference to the overall chronological picture. (See the discussion in sections 12.1 and 14.8 of two difficult areas that particularly affect loanwords from Latin and French, and in the latter case also loanwords from Greek and German.)

Among the 92,500 main entries that make up the alphabetical ranges M, N, O, P, Q, R, and A to ALZ in *OED3*, a little over 29,300 (approximately 32%) are identified as loanwords from other languages. Fig. 2.1 gives the twenty-five most prolific sources (twenty-four different languages, with French and/or Latin counted as a separate category).

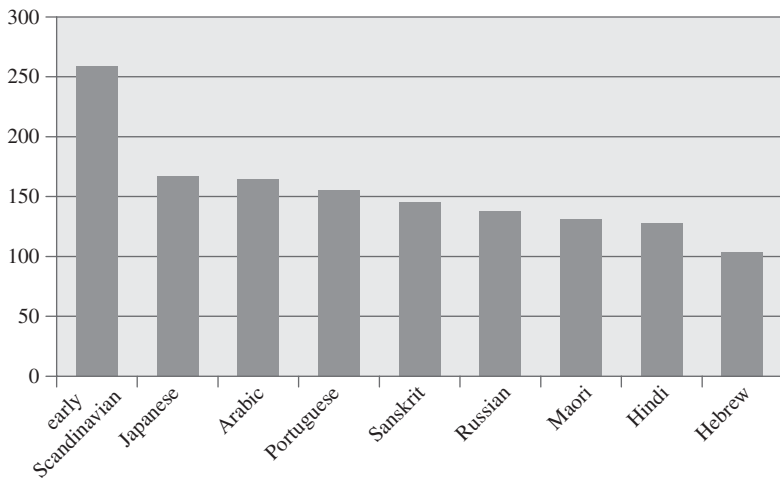
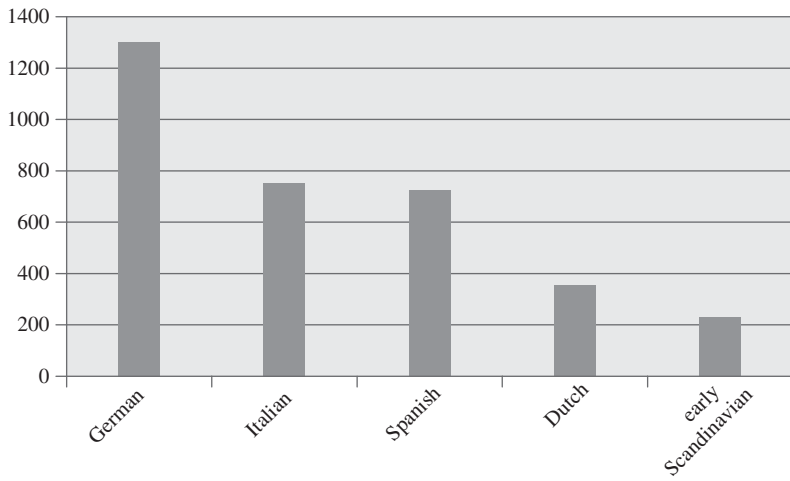
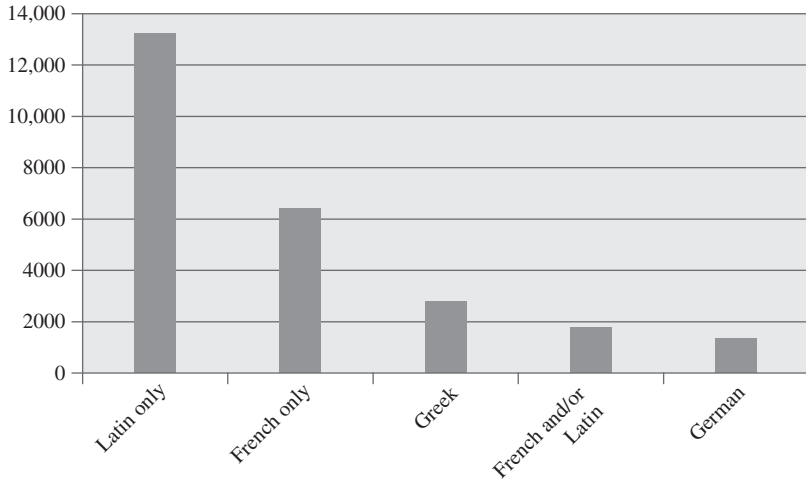
Fig. 2.1 shows how dramatically the totals of loanwords from different languages differ. Those from the top few sources dramatically outnumber those from any other source, such that the totals cannot be made out clearly from a single chart. Fig. 2.2 therefore consists of a series of charts, each to a different scale, enabling the totals to be seen more clearly: in each successive

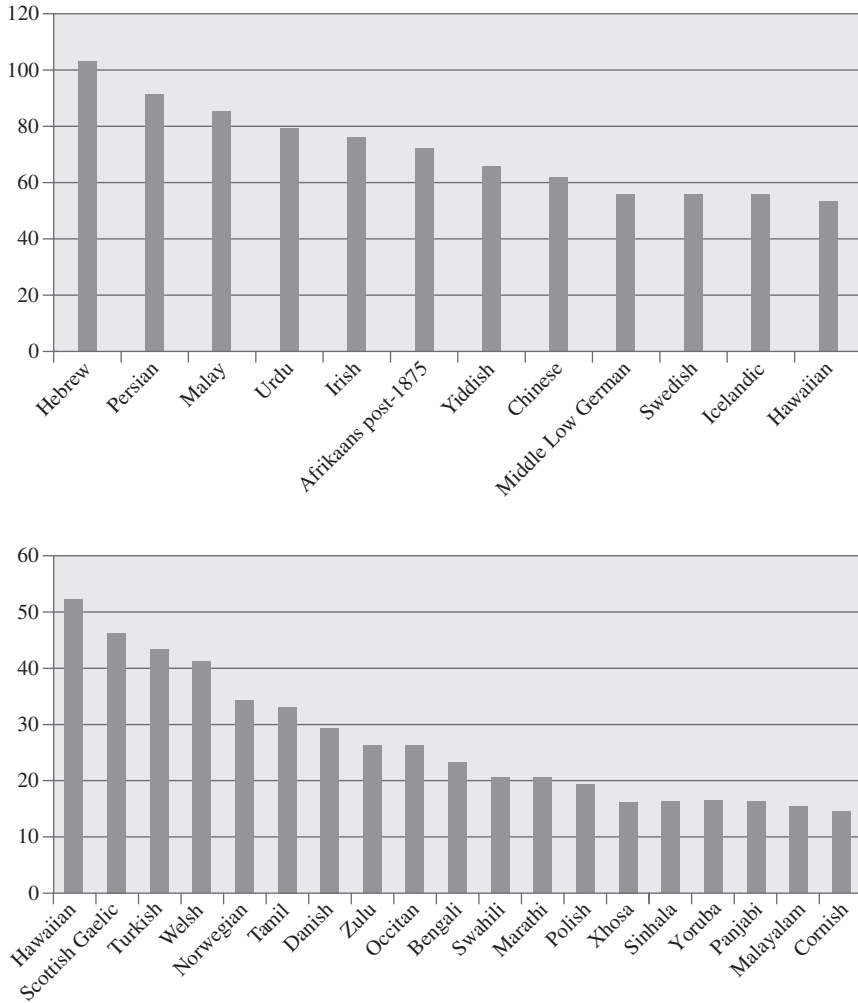
publicly available, and for this reason I have not specified the precise details of the computer searches made. However, all of the relevant data is present in the *OED* entries as published in *OED Online*, and in the companion website to this book I suggest a number of approaches for exploring *OED* data directly using *OED Online*. It should also be noted that newly revised and newly added entries are being added to *OED3* and published on *OED Online* (see chapter 17) every three months, and hence the *OED3* data presented in this book represents a snapshot of *OED*'s research at a particular point in time. See again the companion website for some ways of exploring and engaging creatively with this dynamic aspect of the *OED* today.

<sup>2</sup> The *OED3* text also incorporates all loanwords that have appeared in previous editions of the *OED* (including its one-volume supplement of 1933). No words added by previous editors are omitted, even if the supporting evidence appears slender; the only exception is where evidence appears to have been misinterpreted, giving rise to a 'ghost word'.



**Fig. 2.1** Totals of loanwords from the 25 most prolific inputs in *OED3*.





**Fig. 2.2** The most frequent donor languages, arranged in five charts graded by scale.

chart, the language that was on the right-hand edge of the preceding chart appears first at the left-hand edge. These charts take the totals beyond the top twenty-five donors, so as to include Welsh and Cornish, as long-standing Celtic neighbours of English in the British Isles. (On the Celtic languages see further section 3.3.)

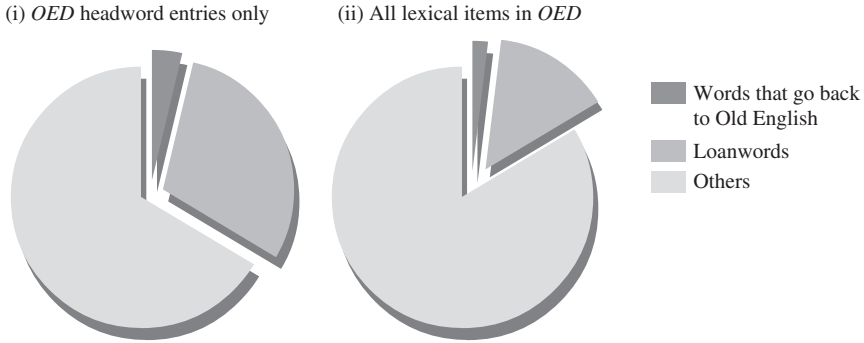
Some of the language names given here may be more familiar than others. See section 15.2.1 for explanation of the distinction between Low German and High German. See section 9.2 for an explanation of what is meant in this book (as in *OED3*) by the cover term ‘early Scandinavian’.

It should be noted that some of the totals in these charts overlap, because many words show input from more than one language, or the language of direct origin cannot be ascertained with precision. For example, many words that may have been borrowed into Middle English or Early Modern English from Low German could alternatively have been borrowed from Dutch, or may show some input from both languages. Some more recent borrowings from Dutch show some input from Dutch as spoken in Europe, some from Dutch as spoken in South Africa, and some from Afrikaans, which developed from this variety of Dutch. (See section 15.2.1.)

By far the biggest group of words of either multiple or uncertain origin is constituted by words from French and/or Latin. In this book, I show that such words entered English in particularly large numbers at a key turning point in the history of the English lexicon, and argue that they are symptomatic of very significant historical changes. In Fig. 2.1 and Fig. 2.2, they are therefore presented as a separate category.

Loanwords from all of the twenty-five most prolific inputs listed in Fig. 2.1 are looked at in the course of this book. However, it should be noted that these raw numbers (in Fig. 2.1 and Fig. 2.2) do not reflect qualitative differences in impact on the vocabulary of modern English. We will look at this topic more closely in sections 2.2 and 2.3. It will be noted that, for example, in these charts the total of loanwords from Japanese comes next after the total from early Scandinavian. However, Japanese has (so far at least) contributed no words to English which form part of the basic vocabulary of modern English, whereas early Scandinavian contributed several, including even the third-person plural personal pronoun *they*. Such qualitative differences are a major determining factor in the amount of discussion devoted to loanwords from various source languages in the course of this book. Additionally, it should be noted that not all loanwords from a particular source language are of similar importance in the vocabulary of English. For instance, the majority of the words that figure in the huge total of English loanwords from Latin are of relatively rare occurrence, but some occur very frequently, and have an important role in the language of everyday discourse. Such words are looked at particularly closely in this book.

In fact, the ultimate contribution of loanwords to the vocabulary of English is much higher than even these numerical totals suggest, if we also take into consideration all of the words formed within English from borrowed words. The inherited portion of the vocabulary of English that



**Fig. 2.3** All items in *OED* classified by origin. (i) *OED* headword entries only. (ii) All lexical items in *OED*.

goes back to Old English is relatively small: the *OED* lists among its 275,000 headword entries approximately 9,600 words (around 3.5%) surviving beyond 1150 that were first attested before 1150. (Even this total includes several hundred pre-1150 loanwords from Latin and smaller numbers of loanwords from Scandinavian and Celtic languages, but these are very approximately counterbalanced by words not recorded until after 1150 that are nonetheless generally considered to be of native origin and to have existed already in Old English.)

Thus, if we put together all of the words found before 1150 and all of the words known to have been borrowed subsequently from other languages, we have only about one-third of the *OED* headword list. Some of the remaining words are of unknown origin, and may be borrowings, while others can be identified pretty confidently as imitative formations reflecting non-linguistic sounds in the external world (e.g. bangs or crashes, animal cries, or natural utterances such as groans), and others are derived from proper names. However, the vast majority are formed from those words that are either inherited or borrowed, chiefly by processes of compounding, derivation, or conversion, e.g. from *fork*, we find *fork handle*, *forked*, and *to fork*. In some cases the relationship with the parent word(s) remains transparent; in others it has become opaque, e.g. in *lord* or *lady*, both of which originally showed compounds of *loaf*.

Fig. 2.3 shows the proportions of words of each type in modern English, as reflected by the *OED*. Here, (i) shows the relative proportions among *OED* headword entries only, while (ii) shows what we find if we take into account all lexical items in *OED*, including the huge numbers of compound

**Table 2.1** Proportions of words of different origins in modern English.

	SOED (%)	ALD (%)	GSL (%)
Inherited <sup>3</sup>	22.2	27.43	47.08
French	28.37	35.89	38.2
Other Romance languages	1.86	1.6	0.2
Latin	28.29	22.05	9.59
Greek	5.32	1.59	0.25
Scandinavian element	2.16	2.51	3.11
Dutch, Low German, Frisian	1.42	1.61	0.7
High German (incl. Yiddish)	0.5	0.28	0
Celtic	0.43	0.32	0.025
Other European languages	0.13	0.11	0
Non-European languages	2	1.12	0.05
Unknown etymology	4.03	3.84	0.98
Proper names	3.29	1.96	0

Source: Scheler (1977) 72

and derivative words that are listed under a parent word as subordinate entries, rather than as separate entries.

Thus, in both of these charts the category ‘others’ is by far the largest. The vast majority of the words in this category are formed by compounding, conversion, or derivation, either from words that go back to the earliest stage of the English language, or from words that have been borrowed from other languages. Attempting to estimate what proportion of the vocabulary of English is formed ultimately from inherited words, and what proportion from words of Latin origin, French origin, etc., is fraught with difficulties. Any figures arrived at can only be very approximate, depending on how one deals with a great many variables. However, if they are approached with due caution, such figures can be illuminating. Scheler (1977 72) presents the figures reproduced in Table 2.1, drawn from Finkenstaedt et al. (1973). The first column shows data taken ultimately from the *Shorter Oxford English Dictionary* (SOED), the second from the *Advanced Learner’s Dictionary of Current English* (ALD, 1963 edition), and the third from a basic vocabulary list, the *General Service List of English Words* (GSL). Broadly the same data from Finkenstaedt et al. (1973) is drawn upon in many of the accounts found in current handbooks of English (often via the analysis in Wermser (1976)).

<sup>3</sup> Including some words of unknown etymology recorded for the first time in Old English or Middle English.

Some important caveats need to be noted about these figures. The *SOED* figures are derived from the first edition of that dictionary, which was based on the first edition of the *OED*. The *SOED* omitted some rarer words, especially obsolete ones, and summarized and in some cases simplified etymologies. This data was then classified under various summary types in Finkenstaedt (1973), including some rather arbitrary decisions about how to classify words compounded from elements of different origins.<sup>4</sup> The *ALD* figures use *SOED*'s etymologies, and only include those words from the wordlist of the 1963 edition of *ALD* that were also found in *SOED*. The *GSL* figures include only its headwords (which are all high-frequency items), rather than subordinate entries, with etymologies being taken from the *Oxford Dictionary of English Etymology (ODEE)*.

One important point to note is that the figure for words from Latin reflects some assumptions on which thinking has changed considerably since the first edition of the *OED* (and since *ODEE*, which mirrored its approach closely): in *OED3*, many of these words are given as potentially being borrowed immediately from French, either partly or wholly (see further section 11.3); additionally, a great deal of scientific vocabulary is classified as directly from Latin (or Greek) in the first edition of *OED*, which in *OED3* is presented as formed within English or in other modern languages from elements ultimately of Latin or Greek origin (see further section 14.8.1).

The figures therefore do need to be approached with caution, from various points of view. However, they are valuable for the perspective they give on the composition of the vocabulary of modern English. It is evident at a glance that the proportion of words ultimately reflecting borrowing from most languages (or language groups) declines steadily as one moves from the fuller wordlist of *SOED* to the high-frequency wordlist of the *GSL*. The exceptions to this are French and Scandinavian, the proportions of which actually increase from left to right of Table 2.1. This reflects some important characteristics of borrowing from these two languages, which are explored in detail in this book.

In the following two sections, I turn aside from the wordlists of dictionaries, to see what light can be shed by corpus data (see section 2.2) and basic meaning lists (see section 2.3) on the contribution of borrowing from

<sup>4</sup> See further Durkin (2002a) on what is lost through this approach.