



The Routledge Handbook of Vocabulary Studies

Edited by Stuart Webb

The Routledge Handbook of Vocabulary Studies

The Routledge Handbook of Vocabulary Studies provides a cutting-edge survey of current scholarship in this area. Divided into four sections, which cover understanding vocabulary; approaches to teaching and learning vocabulary; measuring knowledge of vocabulary; and key issues in teaching, researching, and measuring vocabulary, this *Handbook*:

- brings together a wide range of approaches to learning words to provide clarity on how best vocabulary might be taught and learned;
- provides a comprehensive discussion of the key issues and challenges in vocabulary studies, with research taken from the past 40 years;
- includes chapters on both formulaic language as well as single-word items;
- features original contributions from a range of internationally renowned scholars as well as academics at the forefront of innovative research.

The Routledge Handbook of Vocabulary Studies is an essential text for those interested in teaching, learning, and researching vocabulary.

Stuart Webb is Professor of Applied Linguistics at the University of Western Ontario, Canada. His research interests include vocabulary studies, extensive reading and listening, and language learning through watching television.

Routledge Handbooks in Linguistics

Routledge Handbooks in Linguistics provide overviews of a whole subject area or sub-discipline in linguistics, and survey the state of the discipline including emerging and cutting-edge areas. Edited by leading scholars, these volumes include contributions from key academics from around the world and are essential reading for both advanced undergraduate and postgraduate students.

The Routledge Handbook of Language and Humor

Edited by Salvatore Attardo

The Routledge Handbook of Language and Dialogue

Edited by Edda Weigand

The Routledge Handbook of Language and Politics

Edited by Ruth Wodak and Bernhard Forchtner

The Routledge Handbook of Language and Media

Edited by Daniel Perrin and Colleen Cotter

The Routledge Handbook of Ecolinguistics

Edited by Alwin F. Fill and Hermine Penz

The Routledge Handbook of Lexicography

Edited by Pedro A. Fuertes-Olivera

The Routledge Handbook of Discourse Processes, Second Edition

Edited by Michael F. Schober, David N. Rapp, and M. Anne Britt

The Routledge Handbook of Phonetics

Edited by William F. Katz and Peter F. Assmann

The Routledge Handbook of Vocabulary Studies

Edited by Stuart Webb

Further titles in this series can be found online at www.routledge.com/series/RHIL

The Routledge Handbook of Vocabulary Studies

Edited by Stuart Webb

First published 2020
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
52 Vanderbilt Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2020 selection and editorial matter, Stuart Webb; individual chapters, the contributors

The right of Stuart Webb to be identified as the author of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this book has been requested

ISBN: 978-1-138-73572-9 (hbk)

ISBN: 978-0-429-29158-6 (ebk)

Typeset in Times New Roman
by Apex CoVantage, LLC

For Machiko, without whose support this volume would not have been possible.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Contents

<i>List of Figures</i>	<i>xi</i>
<i>List of Tables</i>	<i>xiii</i>
<i>List of Contributors</i>	<i>xv</i>
<i>Acknowledgments</i>	<i>xx</i>

1 Introduction	1
<i>Stuart Webb</i>	

PART I

Understanding Vocabulary 13

2 The Different Aspects of Vocabulary Knowledge	15
<i>Paul Nation</i>	
3 Classifying and Identifying Formulaic Language	30
<i>David Wood</i>	
4 An Overview of Conceptual Models and Theories of Lexical Representation in the Mental Lexicon	46
<i>Brigitta Dóczi</i>	
5 The Relationship Between Vocabulary Knowledge and Language Proficiency	66
<i>David D. Qian and Linda H. F. Lin</i>	
6 Frequency as a Guide for Vocabulary Usefulness: High-, Mid-, and Low-Frequency Words	81
<i>Laura Vilkaitė-Lozdienė and Norbert Schmitt</i>	
7 Academic Vocabulary	97
<i>Averil Coxhead</i>	

8	Technical Vocabulary <i>Dilin Liu and Lei Lei</i>	111
9	Factors Affecting the Learning of Single-Word Items <i>Elke Peters</i>	125
10	Factors Affecting the Learning of Multiword Items <i>Frank Boers</i>	143
11	Learning Single Words vs. Multiword Items <i>Ana Pellicer-Sánchez</i>	158
12	Processing Single-Word and Multiword Items <i>Kathy Conklin</i>	174
13	L1 and L2 Vocabulary Size and Growth <i>Imma Miralpeix</i>	189
14	How Does Vocabulary Fit Into Theories of Second Language Learning? <i>Judit Kormos</i>	207
PART II		
	Approaches to Teaching and Learning Vocabulary	223
15	Incidental Vocabulary Learning <i>Stuart Webb</i>	225
16	Intentional L2 Vocabulary Learning <i>Seth Lindstromberg</i>	240
17	Approaches to Learning Vocabulary Inside the Classroom <i>Jonathan Newton</i>	255
18	Strategies for Learning Vocabulary <i>Peter Yongqi Gu</i>	271
19	Corpus-Based Word Lists in Second Language Vocabulary Research, Learning, and Teaching <i>Thi Ngoc Yen Dang</i>	288
20	Learning Words With Flash Cards and Word Cards <i>Tatsuya Nakata</i>	304

21	Resources for Learning Single-Word Items <i>Oliver Ballance and Tom Cobb</i>	320
22	Resources for Learning Multiword Items <i>Fanny Meunier</i>	336
23	Evaluating Exercises for Learning Vocabulary <i>Batia Laufer</i>	351
PART III		
	Measuring Knowledge of Vocabulary	369
24	Measuring Depth of Vocabulary Knowledge <i>Akifumi Yanagisawa and Stuart Webb</i>	371
25	Measuring Knowledge of Multiword Items <i>Henrik Gyllstad</i>	387
26	Measuring Vocabulary Learning Progress <i>Benjamin Kremmel</i>	406
27	Measuring the Ability to Learn Words <i>Yosuke Sasao</i>	419
28	Sensitive Measures of Vocabulary Knowledge and Processing: Expanding Nation's Framework <i>Aline Godfroid</i>	433
29	Measuring Lexical Richness <i>Kristopher Kyle</i>	454
PART IV		
	Key Issues in Teaching, Researching, and Measuring Vocabulary	477
30	Key Issues in Teaching Single Words <i>Joe Barcroft</i>	479
31	Key Issues in Teaching Multiword Items <i>Brent Wolter</i>	493
32	Single, but Not Unrelated: Key Issues in Researching Single-Word Items <i>Tessa Spätgens and Rob Schoonen</i>	511

Contents

33	Key Issues in Researching Multiword Items <i>Anna Siyanova-Chanturia and Taha Omidian</i>	529
34	Key Issues in Measuring Vocabulary Knowledge <i>John Read</i>	545
35	Resources for Researching Vocabulary <i>Laurence Anthony</i>	561
	<i>Index</i>	591

Figures

4.1	The hierarchical network model	49
4.2	The semantic feature model	50
4.3	The spreading activation model	51
4.4	The revised spreading activation model	52
4.5	The process of lexical selection	53
4.6	The word-association model	56
4.7	The concept-mediation model	57
4.8	The mixed model	57
4.9	The revised hierarchical model	58
4.10	The conceptual feature model	58
4.11	The shared asymmetrical model	59
4.12	The modified hierarchical model	59
6.1	Vocabulary size and text coverage	86
6.2	Vocabulary size (in lemmas) and cumulative text coverage in the Brown corpus	89
12.1	Panel A demonstrates that the number of occurrences (frequency) to linguistic input generally increases from L2 to L1 for bilinguals and then to monolinguals in their single language. Panel B illustrates asymptotic word processing and demonstrates that low frequency items benefit more from additional occurrences (x to $x + 1$) than high frequency items ($x + 7$ to $x + 8$)	176
20.1	Proportion of correct responses on the immediate and one-week delayed posttests in Nakata and Webb (2016a, Experiment 2)	309
20.2	Learning phase and posttest performance in Nakata (2015a)	310
20.3	Actual vs. predicted scores on the one-week delayed posttest in Nakata and Suzuki (in press)	311
21.1	Bilingualized definition of <i>gum</i> for French readers from WordReference.com	327
21.2	Concordance for word family <i>struggle</i> from a typical school text	331
21.3	Concordance for word family <i>diagnose</i> with more comprehensible contexts first	332
22.1	IDIOM Search results output	342
22.2a	Screenshot from Word and Phrase Info (accessed January 7, 2018)	345
22.2b	More examples of pre-sorted concordance lines	345
25.1	Matrix for deciding what type of MWI knowledge a test/elicitation item is tapping into	389

Figures

25.2	An example test item in the format used in the Word Associates Test	392
25.3	An item from COLLEX	393
25.4	An example test item in the format used in DISCO	393
25.5	Five example items from COLLMATCH	394
25.6	An example test item in the format used in CONTRIX	395
25.7	An example of two test items in the format used in Nguyen and Webb's (2017) collocation test (p. 308)	396
25.8	An example sequence of item presentation in a primed LDT	399
27.1	Example items of the GCT	425
27.2	Example score report (WPLT)	426
29.1	Text segments containing n advanced words	463
30.1	<i>I-squared</i> classification system for different types of vocabulary learning (with both intentional-incidental and isolated-integrated continua)	483
35.1	Citation counts for the search terms "vocabulary" and "corpus"	562
35.2	Screenshot of AntFileConverter converting PDF/docx files to plain text (.txt)	570
35.3	Screenshot of the BootCaT Frontend tool for web-scraping corpus data	571
35.4	Screenshot of AntCorGen in the process of collecting data from the <i>PLOS One</i> open access scientific journal	572
35.5	Screenshot of FireAnt in the processing of collecting data from Twitter under the search term "vocabulary"	573
35.6	Screenshot of RANGE in the processing of five textbook chapters	577
35.7	Screenshot of AntWordProfiler in the processing of five textbook chapters	578
35.8	Screenshot of AntWordProfiler highlighting word levels in a single target corpus	579
35.9	Screenshot of the Compleat Web VP tool for online vocabulary profiling	580
35.10a	Screenshots of the corpus.byu.edu online corpus analysis toolkit	581
35.10b	Main interface to the Corpus of Contemporary American English (COCA) subcorpus of corpus.byu.edu	582
35.11a	Screenshots of the Sketch Engine online corpus analysis toolkit	582
35.11b	A "Word Sketch" result for the search term "vocabulary" in Sketch Engine	583
35.12	Screenshots of the WordSmith Tools offline corpus analysis toolkit	585
35.13	The main interface of AntConc showing a Key-Word-In-Context (KWIC) result for the search term "many" using a small corpus of textbook reading passages	586
35.14	The main interface of TAALES	587

Tables

2.1	What is involved in knowing a word	16
2.2	Homonyms, homographs, and homophones	19
2.3	Conditions affecting vocabulary learning	23
7.1	Top 50 items of the Academic Vocabulary List	102
8.1	A rating scale for finding technical words (as applied to the anatomy text)	115
9.1	Facilitative and difficulty-inducing factors for word learning, processing, and use	132
13.1	Studies estimating vocabulary size in EFL learners from different countries since the 1980s	196
17.1	Priorities for vocabulary learning and teaching inside the classroom	257
17.2	Technique Feature Analysis	261
19.1	The unit of counting in different word lists	291
19.2	Corpora used to develop general service lists of single words	293
19.3	Corpora used to develop general academic word lists	294
24.1	Description of word knowledge	376
26.1	Model research design for investigating vocabulary learning progress after a treatment	409
26.2	Model research design for investigating vocabulary learning progress through different treatments	409
28.1	Nation's (2013) framework of vocabulary knowledge	434
28.2	Expanding on Nation's framework with different measures of vocabulary knowledge	444
29.1	Sample text in three formats: original, tokenized, and typed	455
29.2	Words, lemmas, and families in the example text	456
29.3	Content and function lemmas that occur in the sample sentences	457
29.4	HD-D calculation summary for sample text	460
29.5	Lexical frequency profile for the sample text	462
29.6	Frequencies per million words in the sample text based on the magazine section of COCA	465
29.7	Average concreteness score for words in the sample text	466
29.8	Average lexical decision time (in milliseconds) for words in the sample text	467
29.9	Number of associated words for each word in the sample text	468
32.1	Schematic representation of various classification systems and their correspondence	519
35.1	Common vocabulary research aims	563

Tables

35.2	Summary of characteristic features of corpus resources	564
35.3	Prominent corpora used in vocabulary research	566
35.4	Summary of characteristic features of word lists	574
35.5	Prominent word lists used in (English) vocabulary research	575
35.6	Strengths and weaknesses of online corpus analysis toolkits	581
35.7	Strengths and weaknesses of offline corpus analysis toolkits	584

Contributors

Laurence Anthony is Professor of Applied Linguistics at the Faculty of Science and Engineering, Waseda University, Japan. He is a former director and current program coordinator of the Center for English Language Education in Science and Engineering (CELESE). His main interests are in corpus linguistics, educational technology, and English for Specific Purposes (ESP) program design and teaching methodologies.

Oliver Ballance is a teacher-researcher working on English for Specific Purposes programs at Victoria University of Wellington, New Zealand. Oliver's academic work explores how language technologies can be used to support second language acquisition and development in professional and academic contexts.

Joe Barcroft is Professor of Spanish and Second Language Acquisition and Affiliate Professor of Psychological & Brain Sciences at Washington University in St Louis, USA. His research focuses on second language vocabulary learning, input processing, the bilingual mental lexicon, and psycholinguistic approaches to issues in second language acquisition.

Frank Boers is Professor at the University of Western Ontario, Canada. His early publications were in the field of lexicology, with a focus on polysemy and metaphor. His more recent interests, however, stem from his extensive experience as an EFL teacher. He now publishes mostly about L2 instructional methods.

Tom Cobb pioneered language learning technologies for reading and vocabulary development in many ESL lands in the 1980s and 1990s. He now lives and works in Montreal, Canada, teaching courses at a French-speaking university and conducting research in technology development and applied linguistics. Tom's Compleat Lexical Tutor website (www.lextutor.ca) provides teachers with tools to support language instruction.

Kathy Conklin is Full Professor in Psycholinguistics at the University of Nottingham, UK. A major focus of her research is on the application of psycholinguistic methods to the exploration of lexical and formulaic language processing in a first and second language. Her work has appeared in *Applied Linguistics*, *Brain and Language*, *Bilingualism: Language and Cognition*, and *Second Language Research*.

Averil Coxhead teaches undergraduate and postgraduate courses in applied linguistics and TESOL in the School of Linguistics and Applied Language Studies, Victoria University of

Wellington, Aotearoa/New Zealand. She is the author of *Vocabulary and English for Specific Purposes Research* (2018; Routledge). Her current research includes specialized vocabulary in spoken and written English and other languages.

Thi Ngoc Yen Dang is Lecturer at the University of Leeds, UK. She obtained her PhD from Victoria University of Wellington, New Zealand. Her research interests include vocabulary studies and corpus linguistics. Her articles have been published in *Language Learning*, *English for Specific Purposes*, *Journal of English for Academic Purposes*, and *ITL – International Journal of Applied Linguistics*.

Brigitta Dóczy teaches at the Department of English Applied Linguistics of Eötvös Loránd University in Budapest, Hungary, where she offers courses in vocabulary research, business English, pedagogical grammar, and pedagogical phonology. Her research interests are second language vocabulary development as well as task-based teaching. Recently, she has been doing research comparing the mental lexicon of teacher trainees and translator trainees.

Aline Godfroid is Associate Professor in Second Language Studies at Michigan State University, USA. Her primary research interests are in second language psycholinguistics, the teaching and learning of vocabulary, and quantitative research methodology. She co-directs the Second Language Studies Eye-Tracking Lab at Michigan State University.

Peter Yongqi Gu is Associate Professor at the School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand. His main research interests fall under learner autonomy and learning strategies, vocabulary acquisition, and language assessment.

Henrik Gyllstad is Associate Professor of English Linguistics at the Centre for Languages and Literature, Lund University, Sweden. His research interests include language testing, second language acquisition, and multilingualism, especially to do with vocabulary, phraseology, and lexical processing. His work has appeared in journals such as *Applied Linguistics*, *Language Learning*, and *Studies in Second Language Acquisition*.

Judit Kormos is Professor of Second Language Acquisition at Lancaster University, UK. She has published widely on the topics of individual differences in language learning and the cognitive processes of using and acquiring additional languages. She is also renowned for her research and teacher education activities for teaching languages to students with specific learning differences.

Benjamin Kremmel is a researcher and lecturer at the University of Innsbruck, Austria, where he is heading the Language Testing Research Group Innsbruck. His research interests are in language testing, particularly in the areas of vocabulary assessment, L2 reading assessment, and language assessment literacy. He has published his work in *Language Testing*, *Language Assessment Quarterly*, *Applied Linguistics*, and *TESOL Quarterly*.

Kristopher Kyle is Assistant Professor of Linguistics at the University of Oregon, USA. His research interests include second language acquisition, second language writing, corpus linguistics, computational linguistics, and second language assessment. He is particularly interested in developing natural language processing tools to test/build theories related to

vocabulary development, second language acquisition, and second language writing development (usually using learner corpora).

Batia Laufer is Professor (Emerita) of Applied Linguistics at the University of Haifa, Israel. She has lectured, supervised research, and published widely on several areas of vocabulary acquisition: effective teaching, contribution to reading, testing, cross-linguistic influence, ease and difficulty in learning, dictionary use, and attrition.

Lei Lei is Professor in the School of Foreign Languages, Huazhong University of Science and Technology, China. His research interests are corpus-based studies on vocabulary and academic English. His publications have appeared in *Applied Linguistics*, *International Journal of Corpus Linguistics*, *Journal of English for Academic English*, *Journal of Quantitative Linguistics*, and *RELC Journal*, among others.

Linda H. F. Lin holds a PhD in applied linguistics. She works for the English Language Centre of the Hong Kong Polytechnic University. Her research interests include vocabulary studies, corpus analysis, and applications of concordancing in teaching and learning.

Seth Lindstromberg trained as a secondary school teacher in Canada. However, from 1977 to 2016, he was an EFL teacher in Japan and then the UK, where he also worked as an EFL teacher trainer. His main research interests are pedagogical applications of cognitive linguistics, the semantics of English prepositions, methods for teaching L2 vocabulary, and language teaching methods generally.

Dilin Liu is Professor/Coordinator of Applied Linguistics/TESOL in the English Department at the University of Alabama, USA. His research focuses on corpus-based description/teaching of lexis/grammar. He has published extensively, including numerous articles in many different journals, such as *Applied Linguistics*, *Cognitive Linguistics*, *ESP*, *Journal of English for Academic Purposes*, *Modern Language Journal*, *Studies in the Second Language Acquisition*, and *TESOL Quarterly*.

Fanny Meunier is Professor of English Language, Linguistics, and Didactics at the University of Louvain (UCLouvain), Belgium. She has been involved in learner corpus research for more than 20 years and her main research interest is the link between second language acquisition (SLA) studies and pedagogical applications. She is also actively involved in pre- and in-service teacher training and is collaborating with several international research projects on, among other aspects, bi- and multi-literacies and digital literacies.

Imma Miralpeix obtained her PhD from the University of Barcelona, Spain, where she is a lecturer and researcher. Her main research interests include second language vocabulary acquisition, especially lexical development and assessment, and multilingualism. She is the author of several publications in these areas and has taken part in different funded projects on second language learning and teaching.

Tatsuya Nakata is Associate Professor at Faculty of Letters, Hosei University, Japan. His research interests include second language acquisition and computer-assisted language learning. His research has appeared in publications such as *Studies in Second Language*

Acquisition, Language Teaching Research, International Review of Applied Linguistics in Language Teaching, and The Modern Language Journal.

Paul Nation is Emeritus Professor of Applied Linguistics in the School of Linguistics and Applied Language Studies at Victoria University of Wellington, New Zealand. He is the author of several books and articles on vocabulary and the methodology of language teaching. His website contains numerous free resources for learners, teachers, and researchers.

Jonathan Newton is Associate Professor in the School of Linguistics and Applied Language Studies (LALS), Victoria University of Wellington, New Zealand. His research and scholarship focus on vocabulary teaching and learning, task-based language teaching, and intercultural dimensions of language education. He is passionate about working alongside language teachers to address classroom teaching and learning challenges.

Taha Omidian is a PhD student at Victoria University of Wellington, New Zealand. Taha's research interests include the acquisition, processing, and use of vocabulary and phraseology, academic writing, and learner corpora.

Ana Pellicer-Sánchez is Associate Professor of Applied Linguistics and TESOL at the UCL Institute of Education, UK. Her research centers around the teaching and learning of vocabulary in a second or foreign language, with a particular focus on learning from reading. Her recent research has made use of eye tracking to explore the cognitive processes involved in vocabulary learning.

Elke Peters is Associate Professor at KU Leuven, Belgium. Her research interests involve incidental and deliberate learning of single words and multiword units in a foreign language. She is interested in how different types of input can contribute to vocabulary learning. She has published her research in journals such as *Language Learning, Studies in Second Language Acquisition, Language Teaching Research, and the Modern Language Journal.*

David D. Qian is Professor of Applied Linguistics at the Hong Kong Polytechnic University. He is interested in multifaceted research in applied linguistics, including language assessment, vocabulary studies, and analysis of academic and professional corpora. He co-founded the Asian Association for Language Assessment (AALA) and served as its co-president (2014–2015) and president (2016–2017).

John Read is Professor of Applied Language Studies at the University of Auckland, New Zealand. His research and scholarship have focused on the assessment of second language vocabulary knowledge and the testing of English for academic and occupational purposes.

Yosuke Sasao is Associate Professor of Applied Linguistics at Kyoto University, Japan. His research interests include vocabulary acquisition, language assessment, pedagogical grammar, and academic writing. His articles have been published in journals such as *ITL – International Journal of Applied Linguistics, Language Teaching Research, Language Testing, and RELC Journal.*

Norbert Schmitt is Professor of Applied Linguistics at the University of Nottingham, UK. He is interested in all aspects of second language vocabulary acquisition, use, and assessment,

and has published and presented widely on these topics. He is currently writing a book on language issues for the general public.

Rob Schoonen is Chair Professor of Applied Linguistics at Radboud University, the Netherlands, and Senior Researcher at the Centre of Language Studies, Nijmegen. His research interests are language proficiency, vocabulary knowledge, language assessment, and research methodology.

Anna Siyanova-Chanturia is Senior Lecturer in Applied Linguistics at Victoria University of Wellington, New Zealand. Anna's research interests include psychological aspects of second language acquisition, bilingualism, usage-based approaches to language acquisition, processing and use, vocabulary and multiword items, and quantitative research methods (corpora, eye movements, EEG/ERPs).

Tessa Spätgens completed a PhD at the University of Amsterdam, the Netherlands, on the semantic network and reading comprehension in monolingual and bilingual children. Her linguistic interests include language acquisition, multilingualism, and language teaching.

Laura Vilkaitė-Lozdienė is Assistant Professor at Vilnius University, Lithuania. Her current research focuses on vocabulary acquisition in a second language as well as on identification and processing of formulaic language.

Stuart Webb is Professor of Applied Linguistics at the University of Western Ontario, Canada. His research interests include vocabulary studies, extensive reading and listening, and language learning through watching television. His articles have been published in journals such as *Applied Linguistics* and *Language Learning*.

Brent Wolter is Professor at Ocean University of China and at Idaho State University, USA. His main research area is second language vocabulary acquisition, with a particular interest in how psycholinguistic and cognitive factors affect lexical acquisition and activation. His work has been published in a number of leading journals in the field.

David Wood is Professor of Applied Linguistics at Carleton University in Ottawa, Canada, with a long career in language teaching and applied linguistics research and teaching. He has published and presented extensively on formulaic language in speech fluency and in academic discourse, and also has authored/edited several volumes on formulaic language.

Akifumi Yanagisawa is working on a PhD in applied linguistics at the University of Western Ontario, Canada. His PhD research involves meta-analytic studies of the Involvement Load Hypothesis and Technique Feature Analysis. His research interests focus on second language vocabulary learning and the psychological conditions that contribute to language learning.

Acknowledgments

I am grateful to everyone who contributed to this volume. I know that each person had many commitments, so making the time to write a chapter meant extra work on weekends and nights. However, it is the work of the group as a whole that makes the volume meaningful and worthwhile. Perhaps the best thing about doing research on vocabulary is the opportunity to meet and spend time with the people who work in this area. Not only are the many contributors to this volume great researchers but their kindness, generosity, and spirit make working in this area motivating, interesting, and enjoyable. I believe that the positive culture within the field of vocabulary is the result of the leadership of those who have defined our field for so many years: Paul Nation, Batia Laufer, and Paul Meara. They have been kind, supportive, and generous to so many of us that their influence continues with the next generations of researchers.

I have been very fortunate to have had the chance to learn from and work with a lot of exceptional people. These opportunities have made my contribution to this volume possible. I am most grateful to Paul Nation. Not only is Paul an outstanding scholar, but it is his wisdom, kindness, generosity, and friendship that I find to be most outstanding. I have little doubt that working with Paul was the key to everything good that I have done as an academic. I am also indebted to Frank Boers, whose advice has helped to shape my work and this volume for the better. I would also like to thank my students. Having great students is a joy and I have been very lucky to work with and learn from Thi Ngoc Yen Dang, Yosuke Sasao, Tatsuya Nakata, Michael Rodgers, Akifumi Yanagisawa, Takumi Uchihara, Su Kyung Kim, Zhouhan Jin, and many others. I am also grateful to the many researchers of vocabulary who I have been lucky enough to meet, have a beer or glass of wine with, and consider to be my friends.

Finally, I would like to thank the team at Routledge for their work on the *Handbook* and making its publication a very positive process. I would also like to thank Frank Boers, Ana Pellicer-Sánchez, Elke Peters, Paul Nation, and Takumi Uchihara for their valuable comments and suggestions.

Introduction

Stuart Webb

Introduction

In 1980, Paul Meara's article "Vocabulary Acquisition: A Neglected Aspect of Language Learning" was published. Indeed, at that time there were relatively few studies being published that were focused on the construct of vocabulary, or the teaching and learning of words, and research on second language (L2) vocabulary was particularly scarce. However, perhaps fueled by the influential work of Paul Nation, Batia Laufer, and Paul Meara, there is now much greater interest in studies of vocabulary.

Today, there is a great deal of research conducted on vocabulary, and an increased range in topics relating to words. There are now a number of excellent books that provide a comprehensive review of lexical studies. Nation (2001, 2013) and Schmitt (2010) have provided the widest ranging overviews of vocabulary research in general, while Read's (2000) *Assessing Vocabulary* has been the key book devoted to measuring vocabulary knowledge. More recently, interest in learning sequences of words has generated great interest with Schmitt's (2004) *Formulaic Sequences: Acquisition, Processing, and Use*; Wray's (2002) *Formulaic Language and the Lexicon*; and Boers and Lindstromberg's (2009) *Optimizing a Lexical Approach to Instructed Second Language Acquisition* being among the key texts in this area. Moreover, with the development of new approaches to researching vocabulary, there has also been the publication of books providing guidance to help graduate students and researchers with their research. There are now books on how to develop word lists (Nation, 2016) and how to use eye tracking for data collection (Conklin, Pellicer-Sánchez, & Carrol, 2018; Godfroid, 2019), as well as texts devoted to issues related to researching vocabulary (Nation & Webb, 2011; Schmitt, 2010). Surprisingly, there have been few collective volumes focused on vocabulary. Schmitt and McCarthy's (1997) *Vocabulary: Description, Acquisition and Pedagogy* and Bogaards and Laufer's (2004) *Vocabulary in a Second Language: Selection, Acquisition, and Testing* are the best collections on vocabulary, because each covered a variety of important topics that were written by notable researchers. Recently, there have also been several special issues of journals devoted to vocabulary, including the April 2010 issue of *Reading in a Foreign Language* edited by Averil Coxhead, the May 2012 issue of the *RELC Journal* edited by Paul Nation, the January 2017 issue of *Language Teaching*

Research edited by Batia Laufer, the May 2017 issue of *Language Teaching Research* on multiword units edited by Anna Siyanova-Chanturia, and the April 2018 issue of *ITL – International Journal of Applied Linguistics* edited by Stuart Webb.

The aim of this volume was to build on these earlier texts by providing a comprehensive coverage of vocabulary studies from key researchers in the field today. What makes this volume unique is the wide range of themes on lexis, and the many different perspectives of the researchers. The diversity of issues and foci of the researchers should make it a useful reference text, help to expand on the discussions of vocabulary, and highlight important areas for future research.

Who Is Currently Researching Vocabulary?

One of the goals of the volume was to try to bring different views to some of the subjects. While vocabulary researchers agree on many things, each person brings a different perspective to the research literature. Sometimes these differences are quite small. For example, I believe that many of us (e.g., Averil Coxhead, Jonathan Newton, Irina Elgort, Stuart Webb) who have studied and worked with Paul Nation at Victoria University of Wellington may echo much of what he has written about. There is good reason for this, the biggest being that Paul has written about so many different topics, and even if it takes us a while to recognize it, in the end we usually find that his suggestions are correct. Although we might have differences in our views, these views may often have developed from his earlier work, and so are often similar to his. Moreover, we might consider that there are other schools of thought such as from those who have studied and worked with Paul Meara at Swansea University, and those who have learned from Norbert Schmitt at the University of Nottingham. Paul Meara's influence likely extends to his colleagues, Tess Fitzpatrick and Jim Milton, and the many students who have learned about vocabulary research from them (e.g., Rob Waring, Brent Wolter, Dale Brown, Jeff Stewart). Norbert Schmitt's viewpoint on research has also likely impacted the work of his many students (e.g., Ana Pellicer-Sánchez, Anna Siyanova-Chanturia, Benjamin Kremmel, Laura Vilkaitė-Lozdienė, Beatriz González Fernández) and colleagues (Kathy Conklin, Zoltan Dornyei). Of course all of these researchers have also had a large impact on the wider research community.

With the increase in research on vocabulary there has also been an increase in the number of people investigating it, and these researchers are certainly not limited to those just mentioned. Perhaps nobody has contributed as many seminal studies of L2 vocabulary as Batia Laufer (University of Haifa, Israel). The breadth and influence of her research continues to grow today. Her research initially focused on issues related to lexical difficulty. However, it has touched on a large number of areas related to lexis, such as the development of tests, lexical coverage, formulaic language, and incidental and intentional vocabulary learning. Her work with Jan Hulstijn on the development of the Involvement Load Hypothesis should be considered to be one of the key contributions to vocabulary acquisition research in the last 20 years. While Batia Laufer has focused on a wide range of themes, the impact of John Read in the area of vocabulary testing has had great influence. His book *Assessing Vocabulary* must be considered the essential work in this area, and the continued development and validation of the Word Associates Test should be considered as an important landmark in the development of tests of vocabulary knowledge.

If we think about the different places where people are researching vocabulary, then we can see that there are researchers in a diverse range of countries. With the large body of work completed by Paul Meara, Norbert Schmitt, Ron Carter, and Michael McCarthy in the United Kingdom, not surprisingly there is still a large number of researchers focused

on lexis there, such as Judit Kormos, Dana Gablasova, and Vaclav Brezina (University of Lancaster), Thi Ngoc Yen Dang (University of Leeds), Seth Lindstromberg, Philip Durrant (University of Exeter), and Ana Pellicer-Sánchez (University College London). Elsewhere in Europe, researchers are a little more isolated, with Henrik Gyllstad (Lund University) and Fanny Forsberg Lundell (Stockholm University) in Sweden, Benjamin Kremmel (University of Innsbruck) in Austria, Laura Vilkaitė-Lozdienė (Vilnius University) in Lithuania, Rob Schoonen (Radboud University) and Tessa Spätgens (University of Amsterdam) in the Netherlands, Imma Miralpeix and Raquel Serrano (University of Barcelona) in Spain, and Brigitta Dóczi (Eötvös Loránd University) in Hungary. There is also a lot of interest in vocabulary research in Asia with researchers such as Rob Waring (Notre Dame Seishin University), Atsushi Mizumoto, Tatsuya Nakata, and Alan Hunt (Kansai University), David Beglar (Temple University Japan), Laurence Anthony (Waseda University), Junko Yamashita (Nagoya University), and Yosuke Sasao (Kyoto University) all conducting lexical studies in Japan, whereas, David D. Qian (Hong Kong Polytechnic University) and Anna Chang (Hsing-Wu University) are based in Hong Kong and Taiwan, respectively.

Vocabulary has also been a popular topic among researchers in Belgium, led by the work of Frank Boers. His cognitive linguistics studies on the teaching and learning of single and multiword items have fueled many useful articles by his former colleagues and students: June Eyckmans (Ghent University), Hélène Stengers and Julie Deconinck (Vrije Universiteit Brussel), and Aline Godfroid (now at Michigan State University). Sylviane Granger and her colleagues Fanny Meunier and Magali Paquot at Université Catholique de Louvain have made an impact on research of formulaic language by conducting a large number of corpus-based studies with a particular focus on learner language use. More recently, Elke Peters and Maribel Montero Perez and their students at KU Leuven University have made important contributions to research on vocabulary, perhaps most notably with their research on multimodal learning of words.

In North America, researchers tend to be more scattered. In the US, Joe Barcroft (Washington University in St. Louis), Brent Wolter (Idaho State University), Dilin Liu (University of Alabama), Kristopher Kyle (University of Hawai'i at Manoa), Scott Crossley (Georgia State University), Scott Jarvis (University of Utah), and Dee Gardner and Mark Davies (Brigham Young University) have emerged as key vocabulary researchers. In Canada, Tom Cobb (Université du Québec à Montréal), Marlise Horst (Concordia University), Sima Paribakht (University of Ottawa), David Wood and Michael Rodgers (Carleton University), and the team at the University of Western Ontario (Frank Boers, myself, and students such as Akifumi Yanagisawa, Takumi Uchihara, Zhouhan Jin, Su Kyung Kim) have been frequent contributors to the literature on teaching and learning single and multiword items.

How Might We Classify the Many Investigations of Vocabulary?

With so many lexical studies being conducted, it is perhaps useful to classify research into distinct categories to reveal potential gaps in the literature, as well as to highlight areas that receive more or less attention. There are many different ways in which we might classify research on vocabulary. The following sections are a few of the possibilities.

Description, Pedagogy, and Assessment

This volume is loosely structured around these three categories, as they are at the heart of understanding much of what is written about vocabulary. Description entails understanding what is involved in knowing a word or sequence of words, the factors that influence learning,

how vocabulary learning is related to other components of language development, how we might classify vocabulary, and how vocabulary fits into existing theories of language learning. Description is central to understanding and evaluating research on vocabulary. Nation's (2001, 2013) description of vocabulary knowledge is the most important work on the subject to date, which is why this volume begins with this topic (see Nation, this volume).

Much of the research on vocabulary has focused on pedagogy. Research has examined the efficacy of different vocabulary learning activities (e.g., Laufer & Shmueli, 1997), learners' strategies in vocabulary learning (e.g., Gu & Johnson, 1996), and the tools (e.g., word lists, flash cards, dictionaries) that can be used for teaching and learning words (e.g., Dang, Coxhead, & Webb, 2017; Laufer & Hadar, 1997; Nakata, 2011). A key contribution in this area was Laufer and Hulstijn's (2001) Involvement Load Hypothesis, which sought to identify factors present in text-based activities that contribute to vocabulary learning and provide a testable framework that could predict the relative efficacy of such activities. Before the Involvement Load Hypothesis, the rationale for why more words might be learned through completing one activity rather than another was that in the former the words may have been processed more deeply than in the latter. While this might well be true, it is also a very unsatisfactory statement because there is no means to determine what constitutes deep processing. Although the Involvement Load Hypothesis may also be in need of refinement (e.g., Folse, 2006), it raised awareness of the need to look at the psychological conditions within activities that contribute to learning. Moreover, it also motivated the development of other frameworks such as Barcroft's (2002) TOPRA model and Nation and Webb's (2011) Technique Feature Analysis that shed further light on the relative efficacy of activities (see chapters by Laufer and Lindstromberg, this volume).

Assessment of vocabulary knowledge plays a large role in both pedagogy and research. In the classroom, tests of vocabulary knowledge can motivate study, raise awareness of different aspects of vocabulary knowledge (see Yanagisawa and Webb, this volume), indicate the extent of lexical development within a course (see Kremmel, this volume), and reveal which words students know and which words they need to learn (see chapters by Gyllstad and Kyle, this volume). In research, the quality and focus of tests of vocabulary knowledge may affect whether or not vocabulary learning is found to occur, as well as the number of words participants are found to "know" (see Read, this volume). More sensitive test formats, such as those employing a meaning recognition format, are likely to reveal greater learning or knowledge, while more demanding formats, such as form recall, are likely to show less learning or knowledge (see Laufer & Goldstein, 2004; Godfroid, this volume). Moreover, a mismatch between what is learned (e.g., written form) and what is measured (e.g., form-meaning connection) may provide misleading results (Webb & Piasecki, 2018). There is much that needs to be considered when selecting or designing tests of vocabulary knowledge, and of particular importance is the need to carefully consider the multidimensionality of vocabulary knowledge and assess the aspects of knowledge that are most likely to be learned or known (Nation & Webb, 2011).

Single-Word Items vs. Multiword Items

Although researchers have long been aware of the importance of learning sequences of words as well as individual words (e.g., Palmer, 1933), until recently the vast majority of studies had focused on single-word items. However, with advances in technology, it is now much easier to identify and research multiword items (see chapters by Meunier and Wood, this volume). Indeed, in the last 30 years, more and more has been written about formulaic

language, to the point that there are many of the same lines of research for multiword items as single-word items. For example, there are studies looking at the factors that affect learning multiword items (see Boers, this volume), how formulaic language is processed (see Conklin, this volume), how tests can be created to measure knowledge of multiword items (see Gyllstad, this volume), and the resources that are available for learning formulaic language (see Meunier, this volume).

Although research on multiword items is still in its infancy, there are also some questions that we might hope would be answered by now. What might be the most important question to answer is whether it is more effective to learn words as individual items or as sequences of items, as this may help to optimize instruction. For example, is it better to learn the form and meaning of *take* and encounter it in context, or is it more effective to learn the forms and meanings of the most frequent sequences in which it occurs (*take care*, *take place*, *take advantage*) and encounter each of these in several sentences?

Despite the long history of research on individual words, there are still some very important practical questions that remain unanswered. Perhaps the most important of these questions is how many words can students learn in different periods of time? This is a particularly important question because answering it would allow teachers, learners, and program developers to set meaningful goals for courses, programs, and study periods. If we aim for our students to understand most forms of speech, research suggests that they need to know the 3,000 most frequent word families. A word family is made up of a headword (e.g., *approach*), its inflections (*approaches*, *approaching*, *approached*), and its derivations (*approachable*, *unapproachable*). If the goal is for students to understand written text, then the objective should be that they know the most frequent 8,000 word families, and if we expect them to reach the vocabulary size of an educated L1 user, then they should learn 15,000 word families. Without understanding how many words that students can learn in courses and over different durations, it is unlikely that language learning programs will be highly effective.

Intentional vs. Incidental Word Learning

Intentional and incidental vocabulary learning often seem to be discussed as the only two approaches to learning words. Exercises and activities that are designed to explicitly focus students on learning words are labeled as examples of intentional vocabulary learning, whereas activities that involve learning words through encounters in meaning-focused input are labeled as examples of incidental vocabulary learning (see chapters by Lindstromberg and Webb, this volume). This distinction makes some sense because there are many activities, such as flash cards, fill in the blanks, and matching exercises, that are designed to make students focus on learning words to develop lexical knowledge. There are also many other situations that involve encountering input through reading and listening in which there may be no intention to learn words, and yet through these experiences words are in fact learned.

The advantage of labeling vocabulary learning as intentional and incidental is that it allows us to see the similarities between different learning conditions, as well as their strengths and weaknesses. For example, we might note that vocabulary learning is relatively effective when we learn words intentionally through using flash cards, filling in blanks, and writing words in sentences in comparison to encountering unknown words when reading, listening, or viewing television. The disadvantage of categorizing vocabulary learning as an intentional-incidental dichotomy is that we may lose sight of the many differences among the different intentional learning activities, as well as the differences among the incidental learning activities. Moreover, as with most dichotomies, there may also be the urge to state

that one is good or effective and the other is bad or less effective. However, we should not think of intentional and incidental vocabulary learning as being in competition with each other but rather as useful complements to each other.

Processing vs. Learning Words

Another distinction that is made in research is studies focused on rates of processing words and those that look at the amounts of vocabulary learning (see chapters by Conklin, Godfroid, and Pellicer-Sánchez, this volume). Most studies of vocabulary have focused on word learning perhaps in part because of the ease with which we can assess vocabulary gains using paper and pencil tests. Studies of lexical processing are becoming much more common because improved technology has provided new tools for measuring lexical processing such as reaction time tasks and eye tracking.

Processing speed is sometimes viewed as indicating the strength of links within the mental lexicon, as well as the strength of knowledge that someone has for a word or sequence of words; words that are more closely linked or are better known should be processed more quickly than those that are less closely linked or less well known. In fact, processing speed can also be considered a measure of vocabulary learning because it provides a very sensitive measure of vocabulary knowledge that may be difficult to reveal through more traditional paper and pencil tests. It is probably easiest to see processing speed as a measure of lexical fluency; the faster that we can process a word, the more fluent our access to that word (see chapters by Conklin and Godfroid, this volume).

Depth of Vocabulary Knowledge vs. Breadth of Vocabulary Knowledge

Breadth of vocabulary knowledge refers to the number of words known. In studies of breadth, knowing a word is typically indicated by whether students know the form-meaning connections of words. Depth of vocabulary knowledge refers to how well a word is known, and this is usually indicated by whether students know aspects of knowledge, such as collocation, multiple meanings of words, and derivations, rather than only knowing form-meaning connection. Generally, it is only in the last 20 years that research has started to focus on depth of knowledge. Before that, knowing a word was pretty much always indicated by whether or not students knew form-meaning connection.

Although there has been quite a lot written about depth of vocabulary knowledge, most new studies of words still equate vocabulary learning with gaining knowledge of form-meaning connection. There is good reason to measure knowledge of form-meaning connection. Knowledge of form-meaning connection is essential for comprehension; the more words we understand in a text, the more likely we are to understand that text (Hu & Nation, 2000; Laufer, 1989; Schmitt, Jiang, & Grabe, 2011). Moreover, this is true for not only reading comprehension but also for listening comprehension (van Zeeland & Schmitt, 2013). However, it is also extremely important to recognize the multidimensional nature of vocabulary knowledge; there are many different aspects of word knowledge and so measuring any one aspect of knowledge is only providing a partial evaluation of what might be known. Therefore, while measuring knowledge of form-meaning connection may provide a useful measurement of vocabulary knowledge, it is also a very limited evaluation of what might have been learned (see chapters by Godfroid, Nation, Webb, Yanagisawa and Webb, this volume).

More research on depth of vocabulary knowledge is also needed. Perhaps of greatest benefit would be the development and validation of new measures of vocabulary depth (see

chapters by Godfroid, Gyllstad, Read, Yanagisawa and Webb, this volume). The development of earlier tests, such as the Vocabulary Knowledge Scale (Wesche & Paribakht, 1996) and the Word Associates Test (Read, 1993, 1998), were very important innovations in lexical assessment. However, there is a need to keep improving and expanding on these earlier measures. In particular, it would be useful to have measures of different aspects of vocabulary knowledge such as derivation and collocation with test items selected according to word frequency levels. This would allow teachers and researchers to measure knowledge of the most useful words for different components of depth. Moreover, this would also allow us to compare knowledge of these aspects with the results of established tests that use a similar format such as the Vocabulary Levels Test (Webb, Sasao, & Ballance, 2017).

Quantitative vs. Qualitative Studies of Vocabulary

Within applied linguistics and other research disciplines, studies are often classified by design with the most common classification being quantitative and qualitative studies. Within lexical studies this categorization does not really apply, however, because there are very few qualitative studies of vocabulary (notable exceptions include Gu, 2003; Haastrop & Henriksen, 2000). This is a major limitation of the research on vocabulary. Although we can learn much through quantitative studies of words, there is also a great deal that is unaccounted for that deserves attention. For example, when we look for differences in the amount of word learning through different conditions, we might find that Condition X was more effective than Condition Y. This is useful because it provides some indication of the relative efficacy of the two conditions. However, within each condition, it is likely that there was some variation in the amount of learning; most students may have learned the most through Condition X, but a few may have learned very little. Moreover, a few students may have learned more through Condition Y than through Condition X. Qualitative and mixed methods studies (e.g., Godfroid & Schmidtke, 2013) can help to shed more light on inconsistencies within data sets and expand upon earlier quantitative studies.

Vocabulary and the Four Skills

Another way that we might categorize studies of vocabulary is around the four skills: reading, writing, listening, and speaking. Unfortunately, this would lead to an extremely long section on reading, followed by shorter and shorter sections on writing, listening, and speaking in that order. Most research on vocabulary has looked at the words in written text. For example, we can analyze the vocabulary in corpora of written text to determine word and multiword frequencies and the strength of the relationships between words. We can also look at how reading different types of text contributes to vocabulary learning and how the factors within these texts and the characteristics of the learners affect gains. The reason why most studies have focused on written text is that it is easiest to collect this text type to create corpora, or modify written text to suit research purposes. This has allowed us to learn much about the vocabulary in written text (e.g., Nation, 2006), and the extent to which words can be learned through reading (e.g., Webb, 2008).

There has also been a fair amount of research on vocabulary and writing, with many studies focusing on the lexical richness or lexical variation of learner writing (e.g., Kyle, this volume; Laufer & Nation, 1995). With much improved software, such as TAALES (Kyle, Crossley, & Berger, 2018), AntWordProfiler (Anthony, 2014), and AntConc (Anthony, 2018), that indicates the lexical frequency, formulaic language, and lexical relationships between words in text,

there is a large amount of information that can be revealed about the words used in writing (see chapters by Anthony and Kyle, this volume). One challenge with conducting these studies is that while researchers can analyze the words that participants use in their writing, there is a lack of clarity about the words that are not used. For example, if we are interested in the proportion of lower frequency words that are used in student writing, we cannot know if the figures accurately reflect productive knowledge of words. Instead, the figures reflect the choices that students made about which words to use; some students may choose to focus on accuracy and use more of the higher frequency words that they are most familiar with, others may choose to take more risks and use lower frequency words that they are less familiar with, while others may have written their text according to the frequencies of the words that they knew. Studies of lexical richness in writing do, however, provide very useful indications of the vocabulary that is used between L1 and L2 students (e.g., Crossley & McNamara, 2009).

In contrast to studies of reading and writing, there is relatively little research on vocabulary learning through listening and speaking, or the words and sequences of words encountered in spoken discourse. It is likely that the reason for this is that these are more challenging skills to investigate. For example, it is quite easy to create a corpus of written text that is tens of millions of words in size, because there are billions of words of written text freely available online. However, there is relatively little transcribed spoken text freely available, and so analysis of the vocabulary found in speech is less common. Examining vocabulary learning through speaking and listening is particularly difficult in the classroom, because of the unpredictability of interaction, the large number of variables involved, as well as the challenges of accurately recording the interaction. Despite these challenges, the number of studies investigating the vocabulary of spoken language (e.g., Dang & Webb, 2014), and learning words through speaking (e.g., Newton, 2013; Nguyen & Boers, 2018) and listening (e.g., Peters & Webb, 2018; Vidal, 2003, 2011), appears to be increasing. There would be great value in examining vocabulary learning through speaking and listening in new studies. Hopefully recent studies in this area will reveal useful methodological approaches that can be incorporated into new studies.

Vocabulary and the Four Strands

Nation (2007) introduced the four strands as an approach to L2 curriculum development that would provide diverse opportunities for learning. Although the four strands was initially focused on course or curriculum development, its principles also apply to opportunities for lexical development. The four strands are meaning-focused input, meaning-focused output, language-focused learning, and fluency development. Nation suggests that each strand has similar importance, and that a balance in learning between the four strands should provide the greatest benefit to students.

Meaning-focused input involves learning words incidentally through repeated encounters during reading and listening. Meaning-focused input tasks, such as extensive reading and viewing, focus learners on comprehension rather than vocabulary learning. Through encountering words in meaning-focused tasks, students may learn not only the form-meaning connections of words but also learn how they can be used in context. A large number of studies have investigated incidental vocabulary learning with meaning-focused written input (see Webb, this volume). The greatest potential for further research in this area may be through investigating incidental learning with meaning-focused spoken input.

Meaning-focused output involves developing productive vocabulary knowledge through using words in speech and writing in tasks that focus on communication rather than intentional

word learning. Although there are a reasonable number of studies devoted to vocabulary learning through meaning-focused writing (e.g., Crossley & McNamara, 2009; Laufer & Nation, 1995), there are few studies that have looked at vocabulary learning through meaning-focused speaking. This is an area where further research is clearly needed.

Language-focused learning involves the intentional learning of words through exercises and activities such as sentence production and flash cards. Language-focused learning might be what most students and teachers consider to be at the heart of lexical development, because it involves the deliberate teaching and learning of words. However, when we consider that native speakers know as many as 15,000 to 20,000 word families (Goulden, Nation, & Read, 1990) and that you need to know 8,000 to 9,000 word families to understand English novels and newspapers (Nation, 2006), it should be evident that intentional vocabulary learning on its own will not be successful in helping students reach these targets. There are a large number of studies investigating language-focused vocabulary learning activities (e.g., Webb, 2007, 2009; see also chapters by Laufer and Lindstromberg, this volume). However, the research tends to be limited to examining a relatively small number of different activities (Webb, Yanagisawa, & Uchihara, under review). It would be useful for new studies to look at the extent to which common learning conditions, such as matching exercises, true/false questions, cloze activities, and crosswords, contribute to vocabulary learning.

Fluency development involves processing and using vocabulary at a faster rate. Gains in the rate of processing and using words should be viewed as an indication of vocabulary learning (see Godfroid, this volume). Speed reading, 4/3/2 activity, and repeated reading and viewing are examples of activities designed to promote fluency development. Fluency development has tended to receive less attention than the other three strands, both in the classroom and in research. In recent years there has been increased focus on lexical fluency (e.g., Pellicer-Sánchez, 2015). However, with relatively few studies of fluency development, more research is clearly warranted.

The Organization of This *Handbook*

This volume is organized into four parts: Part I, Understanding Vocabulary; Part II, Approaches to Teaching and Learning Vocabulary; Part III, Measuring Knowledge of Vocabulary; and Part IV, Key Issues in Teaching, Researching, and Measuring Vocabulary. The first part is composed of chapters that are at the heart of researching, learning, teaching, and testing words. These chapters explore what it means to know a word, what affects vocabulary learning, explanation of the different types of words (e.g., academic, technical, high-, mid-, and low-frequency words, formulaic language), lexical processing, vocabulary size, and how vocabulary fits into language learning theories. In order to research, teach, and test words, it is necessary to have a grasp of these topics. The second part is devoted to issues related to teaching and learning words. It begins with the broader areas of incidental and intentional learning and then moves to narrower topics, such as vocabulary learning strategies, word lists, and resources for learning words. All of the chapters in this part are linked with both pedagogy and research; there is typically a large amount of research in each area, and each subject should be considered when developing a program of vocabulary learning.

The third part, Measuring Knowledge of Vocabulary, begins with chapters focused on measuring depth of knowledge of single-word items and measuring knowledge of multiword items. It then moves on to discussion of the different approaches to measuring vocabulary knowledge. A goal of this section was to include an explanation of newer approaches to measuring vocabulary knowledge, such as eye tracking, response times, Coh-Metrix, and

TAALES software, as well as tests that have become recently available, such as the Word Part Levels Test (Sasao & Webb, 2017) and the Guessing from Context Test (Sasao & Webb, 2018). Thus, it should provide a useful update on older books that have focused on the topic of measuring vocabulary, such as Read's (2000) excellent *Assessing Vocabulary*.

The final part, Key Issues in Teaching, Researching, and Measuring Vocabulary, was designed to look at issues that might not have been covered in significant depth in earlier chapters, as well as to provide an overview of important themes from the perspective of other experts in the field. Two chapters focus on research issues related to investigating single-word items and formulaic language. These chapters were included because they are likely to be of interest to graduate students and researchers who are planning to start working in this area. Similarly, two chapters focus on key issues in teaching and learning vocabulary (single-word items are the focus of one chapter, while formulaic language is the focus of the other). These chapters may be starting points for teachers looking to help their students do a better job of vocabulary learning. The final two chapters deal with key issues related to measuring vocabulary knowledge and resources for researching vocabulary. While it is likely that all six of the key issues chapters will touch on some of the topics in the other sections, the overlap should hopefully guide readers to explore different parts of the book in more detail.

When first reading the chapters in this volume, I was often struck by the thought that I would have written things very differently. This variation in discussion of subjects is what I believe to be the greatest value of the *Handbook*. Topics that I felt that I knew quite well were described from different perspectives, and this originality in explanation was informative, interesting, and useful. I was thrilled to have read such a great collection of contributions and I know that this book will be an essential resource for myself. I hope that you will find it equally useful.

References

- Anthony, L. (2014). *AntWordProfiler* (Version 1.4.1) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from www.laurenceanthony.net/software
- Anthony, L. (2018). *AntConc* (Version 3.5.7) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from www.laurenceanthony.net/software
- Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, 52, 323–363.
- Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Basingstoke, UK: Palgrave Macmillan.
- Bogaards, P., & Laufer, B. (Eds.). (2004). *Vocabulary in a second language: Selection, acquisition, and testing*. Amsterdam: John Benjamins Publishing.
- Conklin, K., Pellicer-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge, UK: Cambridge University Press.
- Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing*, 18(2), 119–135.
- Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning*, 67(4), 959–997.
- Dang, T. N. Y., & Webb, S. (2014). The lexical profile of academic spoken English. *English for Specific Purposes*, 33, 66–76.
- Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, 40(2), 273–293.
- Godfroid, A. (2019). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. New York: Routledge.
- Godfroid, A., & Schmidtke, J. (2013). What do eye movements tell us about awareness? A triangulation of eye-movement data, verbal reports and vocabulary learning scores. In J. M. Bergsleithner,

- S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 183–205). Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center.
- Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11(4), 341–363.
- Gu, P. Y. (2003). Fine Brush and Freehand 1: The vocabulary-learning art of two successful Chinese EFL learners. *TESOL Quarterly*, 37(1), 73–104.
- Gu, P. Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46(4), 643–679.
- Haastруп, K., & Henriksen, B. (2000). Vocabulary acquisition: Acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, 10(2), 221–240.
- Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- Kyle, K., Crossley, S. A., & Berger, C. M. (2018). The tool for the automatic analysis of lexical sophistication (TAALES): Version 2.0. *Behavior Research Methods*, 50(3), 1030–1046. <https://doi.org/10.3758/s13428-017-0924-4>
- Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Clevedon, UK: Multilingual Matters.
- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399–436.
- Laufer, B., & Hadar, L. (1997). Assessing the effectiveness of monolingual, bilingual, and “bilingualised” dictionaries in the comprehension and production of new words. *The Modern Language Journal*, 81(2), 189–196.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–322.
- Laufer, B., & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal*, 28(1), 89–108.
- Nakata, T. (2011). Computer-assisted second language vocabulary learning in a paired-associate paradigm: A critical investigation of flashcard software. *Computer Assisted Language Learning*, 24(1), 17–38.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63, 59–82.
- Nation, I. S. P. (2007). The four strands. *Innovation in Language Learning and Teaching*, 1(1), 1–12.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. Amsterdam: John Benjamins.
- Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle Cengage Learning.
- Newton, J. (2013). Incidental vocabulary learning in classroom communication tasks. *Language Teaching Research*, 17(2), 164–187.
- Nguyen, C. D., & Boers, F. (2018). The effect of content retelling on vocabulary uptake from a TED Talk. *TESOL Quarterly*. Early view online version available at <https://doi.org/10.1002/tesq.441>
- Palmer, H. E. (1933). *Second interim report on English collocations*. Tokyo, Japan: Kaitakusha.
- Pellicer-Sánchez, A. (2015). Developing automaticity and speed of lexical access: The effects of incidental and explicit teaching approaches. *Journal of Spanish Language Teaching*, 2(2), 112–126.
- Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through watching a single episode of L2 television. *Studies in Second Language Acquisition*, 40(3), 551–577.

- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10(3), 355–371. <https://doi.org/10.1177/026553229301000308>
- Read, J. (1998). Validating a test to measure depth of vocabulary knowledge. In A. Kunnan (Ed.), *Validation in language assessment* (pp. 41–60). Mahwah, NJ: Erlbaum.
- Read, J. (2000). *Assessing vocabulary*. Cambridge, UK: Cambridge University Press.
- Sasao, Y., & Webb, S. (2017). The word part levels test. *Language Teaching Research*, 21(1), 12–30.
- Sasao, Y., & Webb, S. (2018). The guessing from context test. *ITL – International Journal of Applied Linguistics*, 169(1), 115–141.
- Schmitt, N. (Ed.). (2004). *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, UK: Palgrave Macmillan.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26–43.
- Schmitt, N., & McCarthy, M. (Eds.). (1997). *Vocabulary: Description, acquisition and pedagogy*. Cambridge, UK: Cambridge University Press.
- van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34(4), 457–479.
- Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, 24(1), 56–89.
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61(1), 219–258.
- Webb, S. (2007). Learning word pairs and glossed sentences: The effects of a single sentence on vocabulary knowledge. *Language Teaching Research*, 11(1), 63–81.
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20, 232–245.
- Webb, S. (2009). The effects of receptive and productive learning of word pairs on vocabulary knowledge. *RELC Journal*, 40(3), 360–376.
- Webb, S., & Piasecki, A. (2018). Re-examining the effects of word writing on vocabulary learning. *ITL – International Journal of Applied Linguistics*, 169(1), 72–93.
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL – International Journal of Applied Linguistics*, 168(1), 34–70.
- Webb, S., Yanagisawa, A., & Uchihara, T. (under review). How effective are intentional vocabulary learning activities? A meta-analysis.
- Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus Breadth. *Canadian Modern Language Review*, 53(1), 13–40.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge, UK: Cambridge University Press.

Part I

Understanding Vocabulary



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

The Different Aspects of Vocabulary Knowledge

Paul Nation

Introduction

Knowing a word involves knowledge of a variety of different aspects of knowledge, and these aspects of knowledge can be known to different levels of strength and detail, and to different levels of fluency. The main reason for a teacher to be interested in what is involved in knowing a word is so that the focus and balance of a language course ensures the development of well-rounded, usable vocabulary knowledge. Thus, this chapter focuses on how the various aspects of knowing a word relate to learning, teaching, and testing.

There are several principles that relate to knowing a word.

- 1 Not all aspects of word knowledge are equally important.
- 2 Word knowledge can be described in terms of breadth (aspects), depth (strength), and fluency.
- 3 Word knowledge develops over a period of time.
- 4 Some knowledge is limited to individual words, while other knowledge is systematic.
- 5 Some knowledge needs to be learned, while other knowledge is constructed through common sense and knowledge of the world.
- 6 The difficulty of acquiring knowledge (learning burden) is affected by a variety of factors including regularity of patterning, the learner's L1, other known languages, opportunity and experience, personal commitment, the quality of teaching, and the quality of course design.
- 7 Vocabulary knowledge is most likely to develop if there is a balance of incidental and deliberate appropriate opportunities for learning.
- 8 Learned aspects of word knowledge are affected by a small number of psychological learning conditions.
- 9 Fluency of word knowledge can be a useful learning focus.
- 10 Testing word knowledge requires careful thought about the purpose of testing, the aspects and strength of knowledge to be tested, the effects of test item type, and the people being tested.

Teachers and course designers need to be aware of the various aspects of knowing a word and need to know how to observe and support their development.

Critical Issues and Topics

The most widely known description of what is involved in knowing a word comes from Nation (2013a, p. 49) as shown in Table 2.1.

The receptive-productive distinction runs through each of the nine aspects in this table. Receptive knowledge is the kind of knowledge needed for listening and reading. At its most basic, it involves being able to recall a meaning when meeting a word form. Productive knowledge is the kind of knowledge needed for speaking and writing. At its most basic it involves being able to recall a word form in order to express a meaning. Receptive knowledge is easier to gain than productive knowledge. However, the kind of learning that is done should match the kind of knowledge needed. So, if a learner’s goal is to read the language, then the most effective kinds of learning will involve incidental learning while reading and deliberate receptive learning using flash cards. If however a learner needs to use all the four skills of listening, speaking, reading, and writing, then there needs to be incidental learning through all four skills and both receptive and productive deliberate learning (Griffin & Harley, 1996; Waring, 1997; Webb, 2009). The two principles that lie behind these research findings is that we learn what we focus on, and we should focus on what we need.

Let us now look at each of the parts of Table 2.1 to see the kind of knowledge involved, and how it might be learned.

Table 2.1 What is involved in knowing a word

Form	Spoken	R	What does the word sound like?
		P	How is the word pronounced?
	Written	R	What does the word look like?
		P	How is the word written and spelled?
Meaning	Word parts	R	What parts are recognizable in this word?
		P	What word parts are needed to express the meaning?
	Form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	Concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	Associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	Grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	Collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	Constraints on use (register, frequency, . . .)	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?

Note: In column 3, R = receptive knowledge, P = productive knowledge.

Source: Adapted from Nation, 2013a

Knowing the Spoken Form of Words

One of the early stages in learning the spoken form of a word involves learning any new sounds that are not in the L1, and at a more general level developing awareness of how sounds can fit together, for example in consonant clusters and consonant vowel combinations. Each language has its own collection of sounds and permitted sound combinations, and even young native speakers have a feeling for what combinations are normal and what are not. Some languages have a variety of consonant clusters while others have few or none. For a native speaker, the learning of sounds is largely systematic, with highly contrasting open and closed sounds being learned early (this partly explains why children's words for mother and father are very similar in different unrelated languages), and with a roughly predictable order of learning sound features. With foreign language learners the first language sound system has major positive and negative influences. Age of learning is strongly related to the likelihood of second language learners acquiring a native-like pronunciation, with younger learners more likely to be successful. There are several explanations for this, each emphasizing maturational, cognitive, or affective factors. The maturational explanation suggests around a certain age that there are physical changes in the brain that make the learning of a new sound system difficult. The cognitive explanation suggests that as the first language sound system becomes more strongly established, it becomes a kind of filter that influences a learner's view of a different sound system (Flege, 1981). The affective explanation says that our pronunciation is an important part of our identity and taking on a new pronunciation, even for a different language, is seen as having to change an important part of who we are (Stevick, 1978). Each of these explanations requires a different approach to learning, with some experimental ways of dealing with the affective explanation involving the consumption of alcohol or the use of chemical relaxants (Guiora, Beit-Hallami, Brannon, Dull, & Scovel, 1972; Brannon, & Dull, 1972).

Acquiring a stable pronunciation of words is important for vocabulary learning, especially for young learners (Service & Kohonen, 1995), because one way that words can enter long-term memory is through the phonological loop, and a stable pronunciation is needed for a word to enter the phonological loop. Older learners however can draw on a wider range of memory strategies beyond formal repetition.

Developing knowledge of the spoken forms of words occurs across the four strands of learning through meaning-focused listening to input, learning through having to engage in spoken communication (meaning-focused output), deliberate learning and teaching, and spoken fluency development. Some adult learners of foreign languages stress the importance of knowledge of articulatory phonetics in improving their pronunciation.

Knowing the Written Form of Words

It is possible to learn another language without learning to read or write it, but especially for learners of English as a foreign language, being able to read allows access to a very large amount of graded reading material, which can provide an enormous boost in developing language proficiency.

Learning to read has a strong phonological basis (Perfetti & Lesgold, 1979) and this is especially true in languages whose writing system is systematically related to the spoken language through an alphabetic or syllabic writing system. An early requirement for learning to read an alphabetic language like English is phonological awareness. In essence this is the realization that words can be broken into separate sounds, and separate sounds can combine

to make words. Some children are not ready to deal with this before the age of six while others are ready much earlier. A very useful activity to develop phonological awareness in a young child is to play games like this – What word is this /p/ – /e/ – /t/? (The separate sounds are not the names of the letters as when saying the alphabet, but are the sounds). When the child gets good at doing this, then the child can take the role of breaking the word into sounds to test the listener. After phonological awareness, the next important piece of knowledge in learning to read is the alphabetic principle, that is, that sounds can be represented by letters. In some languages this representation is very regular and predictable. In English, there are many variants and exceptions, but there is still a core of regularity.

Learning the written form of words needs to occur across the four strands of meaning-focused input (learning by reading texts at the right level), meaning-focused output (having to write words and sentences), language-focused learning (deliberately learning letter shapes, sound-spelling correspondences, and word attack skills, and memorizing irregular words), and fluency development (doing plenty of very easy reading).

Knowing Word Parts

For English, knowledge of word parts primarily involves being able to use the inflectional system of the language, with the next step involving the more gradual growth of knowledge of the derivational affixes. Mochizuki and Aizawa (2000) tested Japanese learners' knowledge of English affixes, finding that affix knowledge increased with vocabulary size, and that there were notable gaps in their knowledge. Schmitt and Meara (1997) and McLean (2017) also found a relationship between vocabulary size and affix knowledge. Sasao and Webb (2017) developed a comprehensive Word Part Levels Test which was used with a variety of learners to propose levels of affix knowledge. The better-known affixes tended to be those that occurred more frequently. There were also high correlations between item difficulty estimates for learners with the same L1 and the total participants' item difficulty estimates. This showed that the first language did not play a strong role in knowledge of English affixes, and learning opportunity through input was likely to have had the greatest effect.

Knowledge of word stems (Wei, 2015), such as *pos* (put), *vers* (turn), and *cept* (take), is best not considered as an aspect of knowing a word, as this knowledge is largely meta-cognitive and not obvious to most native speakers. Word stem knowledge is most usefully developed as a mnemonic device for linking form and meaning.

Connecting Form and Meaning

Being able to recognize or produce the spoken or written form of a word is not much use unless the form has a connection with a meaning. In terms of language use, the most important aspects of vocabulary knowledge for a learner of English as a foreign language are knowledge of the word form and the form-meaning connection. This is because in order to start to read and listen, a learner needs to be able to recognize the form of useful words and be able to attach a meaning to them. For beginners this meaning will be an L1 word. The form-meaning connection is simply attaching a known form to a known meaning.

It is possible to know the form of a word and to know a meaning for a word and yet not realize that the two are connected. This is a bit like the common enough phenomenon of being familiar with someone's name through having heard about them and not realizing that you have already met that person but did not know that that was their name. This phenomenon is common enough in learning a foreign language because in the early and intermediate

stages of foreign language learning, first language concepts, usually in the form of translations, are used as the initial meanings for foreign language words. Thus for many foreign language words, initial learning involves learning a word form and making the form-meaning connection with the first language concept.

It is worthwhile separating out the form-meaning connection as an aspect of knowing a word because there are very helpful learning conditions, such as receptive and productive retrieval, varied retrieval and elaboration, which can strengthen this aspect of knowledge. The very well-researched keyword technique (Pressley, 1977) is a mnemonic technique specifically designed to make the form-meaning connection. Research by Deconinck, Boers, and Eyckmans (2017) shows that getting learners to consider whether the form of a newly met L2 word fits its meaning has very positive effects on establishing the form-meaning connection. This technique is somewhat like the etymological analysis that Boers and colleagues (see, e.g., Boers and Lindstromberg, 2009) advocate for learning multiword figuratives.

The form-meaning connection is usually easy to make for cognates and loanwords. This is because of the closely similar forms; for example, *revolusi* (in Indonesian) and *revolution* (in English) share a roughly similar meaning.

Learning the Concept and Referents

Most words have a core meaning that runs through all or most of their uses. For example, the word *green* has many listed senses in dictionaries referring to color, vegetables, lack of ripeness, inexperience, and so on, but when we look at them, they all seem to share a common core meaning. This should not be surprising because they are all signaled by the same word form. As we might expect, it is easier to learn a new sense for a word than it is to learn a completely new word, especially if the new sense and the known senses have a common underlying meaning (Bogaards, 2001). Thus, it is a useful teaching and learning strategy to draw attention to core meanings and to consider how newly met senses relate to known senses. When using a dictionary for example, it is worth looking at all senses of a word to see what is common, rather than just focusing on the relevant contextual sense.

Some words have completely unrelated meanings for the same spoken or written form. Usually these unrelated meanings have different histories and it is an accident that they share the same word form. Table 2.2 provides the technical terms for the types of relationships with some examples.

The most frequent 2,000 words of English (West, 1953) contains seven homographs (*close*, *lead*, *minute*, *present*, *row*, *wind*, and *wound*), 55 homonyms, and 147 homophones (Parent, 2012). The *Academic Word List* (Coxhead, 2000) contains 60 homonyms and homographs, but in only five cases would both words meet the criteria for inclusion in the list – *issue*, *volume*, *objective*, *abstract*, and *attribute* (Wang & Nation, 2004).

A feature of homonyms and homographs is that it is very unusual for two words like *bear* (to carry) and *bear* (the animal) to have roughly similar frequencies of occurrence.

Table 2.2 Homonyms, homographs, and homophones

	<i>Spoken form</i>	<i>Written form</i>	<i>Meaning</i>	<i>Examples</i>
Homonyms	The same	The same	Different	band (group) – (hoop/ring)
Homographs	Different	The same	Different	minute (time) – (very small)
Homophones	The same	Different	Different	peace – piece

Typically one word makes up well over 70% of the total occurrences. *Bear* (the verb) for example accounts for 92% of the occurrences of the form and *bear* (the animal) 8%. In the most frequent 2,000 words of English only six homonyms (*bowl, ring, rest, net, yard, miss*) have roughly similar frequencies (see Nation, 2016, Chapter 3 for more detail).

Where possible, homonyms, homographs, and homophones should not be taught together, and the most frequent items should be taught first.

Developing Associations With Words

Associations between words (Miller & Fellbaum, 1991) are largely developed incidentally through receptive and productive language use. There are also associations which occur through knowledge of the world and through common sense. There are some conventional associations, such as opposites, synonyms, and hyponyms, that can be established through deliberate learning, but there is likely to be little value in teaching them.

There are numerous activities that involve classifying words into groups and arranging them into semantic maps which can be seen as a form of elaboration which may help strengthen learning.

Learning the Grammar of Words

Part of word knowledge involves being able to use words. Some of this knowledge relates to language systems and some relates to particular words. Let us look at systematic knowledge first. English nouns may be countable or uncountable and this affects whether they can have singular and plural forms, whether they can be used with numerals, articles, and determiners like *much, many, each, and every*, and subject-verb agreement. English verbs can be transitive and intransitive, which affects the use of the passive and what can occur after the verb. Some adjectives can be modified for degree, some taking *-er* and *-est*, and others *more* and *most*.

For many learners of English as a foreign language, learning this kind of systematic knowledge involves profound conceptual development, particularly if the particular knowledge is not similar to first language use. Learning the singular plural distinction, for example, is much more difficult than learning to add *-s* for plurals and so on. It involves developing the concepts of countability and uncountability and applying this view of nouns to their classification and use. It is not surprising that even very advanced learners of English as a foreign language often struggle with singular and plural.

There is some value in giving deliberate attention to grammar, especially for consciousness raising and for self-monitoring of language production. However it is likely that the systematic knowledge of grammatical features which underlies normal language use is ultimately the result of incidental learning from large amounts of meaningful comprehensible input. It is thus useful to see the learning of grammar occurring across the four strands of meaning-focused input, meaning-focused output, language-focused learning, and fluency development. Three of these strands involve incidental learning and one involves deliberate learning. This is probably about the right balance for learning grammar. Ellis (2005) describes an excellent set of principles for the learning of grammar which largely fits with the four strands.

Some grammar knowledge is word-based. That is, it relates to particular words. It is at this point that grammatical knowledge and collocational knowledge overlap. Knowing the word *agree* involves knowing that it is used as a verb, that it is intransitive but it can occur in

sentences such as *We are all agreed that . . .*, that when it is not a simple sentence (*I agree*) it is typically followed by a preposition group beginning with *with*, or an object clause beginning with *that*. No other verb takes the same set of patterns. When beginning to learn the word *agree*, it makes sense to memorize one or two of its most frequent uses in example sentences. This memorization should include analysis and understanding of the parts of the example sentences, rather than simply memorizing them as unanalyzed wholes. This can also be done through the use of substitution tables.

Concordance analysis can be a useful way of gathering information for such teaching and for intermediate and advanced learners the use of concordances can support learning (Cobb, 1997).

Dealing With Collocations

Sinclair (2004) puts the case most strongly for giving attention to collocations: “The lexical unit is best described maximally, not minimally” (p. 81). However, this should be taken to mean that in addition to giving attention to words as units we should also see how they behave in larger units. This is because the vast majority of multiword units are made up of parts where the meanings of the parts make an obvious contribution to the meaning of the whole, and where the parts behave grammatically and semantically in ways that are consistent with their use in other places (Liu, 2010). That is, collocations are not arbitrary groupings of words but are typically regular predictable combinations.

The biggest problem in dealing with research on collocation is coming up with a definition of what can be considered a collocation and then following that consistently. For the purposes of language learning, it is useful to classify collocations into core idioms, figuratives, and literals (Grant & Nation, 2006). Core idioms make up a small number of collocations in English, probably around 100. In core idioms, the meanings of the parts do not clearly relate to the meaning of the whole. It is likely that early in their history there was a connection but this is now not known. The most common core idioms include *as well (as)*, *by and large*, *out of hand*, *serve someone right*, and *take someone to task*.

Figuratives have two related meanings – a literal meaning and a figurative meaning. Here are some examples – *saved by the bell*, *give someone the green light*, *walking on air*, *between a rock and a hard place*. Many figuratives are related to a particular topic area – *out for the count* relates to boxing, *toe the line* relates to military drill, and *threw in his hand* relates to card playing. Gaining awareness of the origin of a figurative helps learning (Boers & Lindstromberg, 2009).

The meaning of literals is closely connected to the meanings of the parts, although it is possible to set up a scale of transparency from those very clearly connected (*like ice cream*) to those containing elements of opaqueness (*put up with*, *about time*). Martinez and Schmitt (2012) have a very carefully made lists of the most useful of these semi-opaque collocations.

These three meaning-transparency-based categories of collocations each have their different language-focused learning approaches (Grant & Nation, 2006). Core idioms need to be memorized as complete units. Because most of them are not frozen but can have different forms (*you’re pulling my leg*; *your leg’s being pulled*; *pull the other one*, *it’s got bells on it*), it is worth giving some attention to the parts. The learning of some core idioms may be helped by creating false etymologies, such as *cats and dogs* in the expression *raining cats and dogs*. Figuratives need to be dealt with using the obvious strategy of relating the literal meaning to the figurative meaning, and where possible finding out the topic area of the literal meaning. Literals can be largely learned incidentally, although in the early stages of learning

a language it is well worth memorizing very useful expressions (*How much does that cost?, Thanks very much, I'd like . . .*) in order to quickly develop spoken fluency (Nation & Crabbe, 1991). Semi-opaque literals require a mixture of analysis and memorization.

The fluency development strand of a course sets up good conditions for collocational knowledge to grow and strengthen. There are several reasons for this. Firstly, receptive fluency development activities (listening and reading) should involve large amounts of input and this can increase noticing and repetition. Secondly, fluency activities often involve some pressure to go faster, and this pressure can encourage learners to restructure their knowledge to work with larger units of language, with words rather than letters, and with phrases rather than words (McLaughlin, 1990). Thirdly, repetition in input usually involves varied repetition rather than verbatim repetition (Webb & Nation, 2017), and varied retrieval is a very effective learning condition (Joe, 1998). The fluency development strand and the other incidental learning strands of meaning-focused input and meaning-focused output are important for all aspects of vocabulary knowledge. Language-focused learning, of which teaching is only a part, needs to be accompanied by learning through use.

A lot of work on collocations has focused on formulaic sequences (Schmitt, 2004; Wray, 2008). Formulaic sequences are units of language that are most likely stored as whole units for the purposes of language use. The motivations for such storage are frequency of occurrence (If a phrase or sentence is used or met often, it is more efficiently stored as a whole unit) and irregularity (If a phrase or sentence cannot be easily reconstructed or analyzed, then it needs to be treated as a unit). Such storage contributes to fluency.

Managing Constraints on Use

A rather small number of words and phrases are marked by restrictions on their use, and being aware of these restrictions is one aspect of knowing a word. Swear words are the most striking examples of such restrictions. There are numerous situations where they cannot be appropriately used. Other restrictions on use include politeness restrictions (words like *fat, old, stupid* need to be avoided when talking to the person they apply to), geographical restrictions (dialects such as US English vs. UK English), age restrictions (language used to talk to children), datedness restrictions in that some words are old-fashioned, and frequency restrictions in that some words are so rare that they sound strange when they are used for common situations.

A very wide ranging restriction in English comes from the very important contrast between spoken and written language (Biber & Conrad, 2009). In formal writing the use of colloquial spoken expressions seems inappropriate. In friendly spoken language and in friendly letters and emails, formal written language sounds unfriendly. In English, this contrast comes partly from the Germanic vs. Latinate vocabulary contrast, with Latinate vocabulary being more formal. Learning about restrictions on use can occur through feedback on use, and through informed observation of proficient users of the language. A few words, like swear words, need to come with a health warning when they are first learned.

The Conditions for Learning

Listing the aspects of what is involved in knowing a word makes vocabulary learning and vocabulary teaching seem like formidable tasks. This is of course partly true. Learners need to know a lot of words and there is a lot to know about each word. However, let us now look at how these formidable tasks are made manageable, but before that it is useful to briefly

Table 2.3 Conditions affecting vocabulary learning

Number of meetings	Initial occurrence/repetition	
	Incidental attention	Deliberate attention
Quality of attention		
Noticing		
Receptive or productive retrieval		
Varied receptive meetings or productive use		
Elaboration		

Source: Adapted from Webb & Nation, 2017

consider the psychological conditions that favor learning. Table 2.3 summarizes these conditions (Webb & Nation, 2017).

The two major conditions in Table 2.3 are *Number of meetings* and *Quality of attention*. Essentially, vocabulary learning depends on how often words are met and the quality or depth of the mental processing at each meeting (Nakata, 2011). The major contrast regarding quality of processing is between incidental attention and deliberate attention, with deliberate attention typically resulting in stronger learning. Both incidental and deliberate attention can occur with different degrees of quality. Noticing is the most superficial but nonetheless still useful level of attention. For incidental learning this could involve simply noticing an unknown word during reading or listening. For deliberate learning this could involve looking up the word in a dictionary or making a word card containing the word and its meaning. The next level of attention involves retrieval and depends on previous noticing. Receptive retrieval occurs when a learner sees or hears a word and has to recall its meaning. This can occur incidentally while reading or deliberately when working with flash cards. Productive retrieval occurs when a learner wants to express a meaning and has to recall the appropriate spoken or written word form. Receptive and productive retrieval are more effective for learning if they involve some degree of difference from previous retrievals or previous noticings, that is, varied meetings or use (Joe, 1998). Fortunately, when words are met again incidentally during listening and reading, they typically occur in different contexts. Elaboration involves some enrichment during the meeting with a word. In incidental learning, this enrichment can come from meeting or having to produce a word in a memorable communicative situation. During reading, a memorable picture may result in elaboration by enriching knowledge of the word, or reading the word on a sign or label may make its occurrence memorable. In deliberate learning, elaboration can occur through the use of a mnemonic technique such as the keyword technique, through word part analysis, or through the analysis of core meaning in a dictionary entry.

The receptive-productive distinction runs through all the learning conditions in Table 2.3, including number of meetings. In general, productive meetings are more demanding and more likely to result in stronger learning than receptive meetings (Griffin & Harley, 1996). The ranking of the levels of quality of attention is largely for explanatory purposes and is definitely not to suggest that ideally all meetings should involve deliberate productive elaboration. However, there are small but effective ways of increasing the quality of attention that learners and teachers could use. Here are some of the most useful ways.

- 1 Before looking up a word in a dictionary try to guess or recall its meaning. This replaces noticing with retrieval or elaboration.

- 2 Use flash cards rather than vocabulary notebooks. Notebooks typically present the form and meaning together (noticing) rather than encouraging retrieval of the form or meaning (Nakata, 2011).
- 3 Encourage extensive reading of graded readers (Day & Bamford, 1998). This provides large amounts of repetition of vocabulary and involves varied receptive retrieval. Talking about what has been read provides an opportunity for productive varied retrieval.
- 4 Use linked skills activities (Nation, 2013b, Chapter 15) where learners deal with the same material across three different skills, for example, they may read a text, talk about it, and then write about it. This encourages repetition, retrieval, and varied meanings and use.
- 5 At the beginning and end of a class, get the learners to recall what they covered in previous classes or in the present class session. This encourages deliberate attention, retrieval, and perhaps elaboration. It can also help move receptive learning to productive use.

Although we have mainly looked at learning word form and the form-meaning connection in the preceding examples, the conditions for learning apply to all aspects of knowing a word. Let us now look at how the aspects of knowing a word are likely to be learned.

Developing Word Knowledge

As has been noted several times in this chapter, learning is not solely dependent upon teaching, but it occurs across the four strands of meaning-focused input, meaning-focused output, language-focused learning, and fluency development. Teaching makes up part of the language-focused learning strand, sharing time with deliberate study. If there is plenty of input and the chance to produce language under both easy and slightly demanding conditions, then a lot of incidental vocabulary learning will occur.

Vocabulary learning is a cumulative process, both in increasing the number of words known and in increasing depth of knowledge of words (Read, 2004). Each word needs to be met several times in a variety of ways and we should expect knowledge of each word to grow and strengthen over time rather than expect each word to be fully learned on the first meeting. A teacher's concern should not be with how a word should be introduced to the learners, but with how it can be met multiple times in a variety of contexts.

The learning burden of a word is the amount of effort needed to learn it. Words differ in their learning burden, with some words being very easy to learn because they are like L1 words, and others requiring various degrees of effort. The learning burden of a word depends on its relationship with L1 words or with words in other languages that the learner knows, and on its regularity with regard to the systems of form, meaning, and use within the L2. Speakers of European languages which are related to English may find that they already know a lot about English vocabulary because the same words occur within their first language. For example, around 80% of the words in the Academic Word List (Coxhead, 2000) have roughly similar forms and meanings in Spanish. Now, with lots of borrowings of English words into languages such as Japanese (Daulton, 2008), Thai, and Indonesian, learners will already know some English vocabulary even before they begin to learn English.

Some words in a foreign language may have a regular predictable spelling while others have an unusual spelling. The senses of some words stay close to their core meaning. The grammar of some words is largely predictable from their part of speech. Research by De Groot (2006) suggests that words that are easy to learn because they fit into regular systems

are not only easier to learn but are also well retained. That is, high levels of learning effort are not essential for good retention, if what is being learned fits into known patterns.

If teachers have knowledge of the learners' L1 and are familiar with the writing system, morphology, and grammar of the L2, they can readily work out the learning burden of L2 words and direct deliberate attention to the aspects of knowledge that most strongly affect the learning burden of a particular word. For example, words like *one*, *yacht*, and *receive* need attention to spelling. Words like *fork*, *sweet*, and *agree* need attention to core meaning. Words like *discuss*, *enjoy*, and *police* need attention to grammar.

The various aspects of word knowledge are not equally important. For initial learning we would expect spoken word form and the form-meaning connection to be the first aspects that would be learned for most words. This knowledge allows the beginnings of comprehension. For a survival vocabulary (Nation & Crabbe, 1991) intended for productive use, spoken word form, the form-meaning connection, and some very basic grammatical knowledge would be important. Other aspects of knowledge can become focuses of attention as proficiency develops.

Future Directions

The model of word knowledge used in this chapter is not at all sophisticated. It is static and treats the various aspects as unrelated parts. It is a convenient way of covering a range of kinds of knowledge but it does not represent vocabulary in use. It also does not represent how a vocabulary develops. A model more focused on use and growth is likely to provide insights that can enrich learning and research. The work of Paul Meara (2006) has gone a long way towards doing this, and this work needs to be continued.

Knowledge of vocabulary develops in many ways, as a result of deliberate learning, direct teaching, incidental learning, transfer from the L1, knowledge of language systems, and the integration of language knowledge with real-world knowledge. There is a lack of longitudinal studies that consider learning from a variety of sources and examine how word knowledge develops over time and under what conditions. Such studies need not all be long-term but can look at what happens to word knowledge over short periods of time. For example, Barcroft's (2007) study of opportunities for word retrieval while reading has useful implications for glossaries and dictionary use. Similarly, studies of what happens to particular words during extensive reading (Horst, 2005; Pigada & Schmitt, 2006; Pellicer-Sánchez & Schmitt, 2010) have useful messages for the use and design of graded readers. Pellicer-Sánchez's (2016) eye-tracking study provided a fascinating view into what can happen to previously unknown words while reading. We need more process-focused studies of this kind and quality.

Research on morphological knowledge (McLean, 2017; Mochizuki & Aizawa, 2000) shows that some groups of learners have very poor knowledge of the derivational affixes of English. This knowledge is way below what learners should know at their current proficiency level. This has a major effect on their vocabulary size and their ability to cope with vocabulary while reading. There is virtually no research which shows how this knowledge can be quickly developed through deliberate learning. There is also a need to examine the role of extensive reading and extensive listening in supporting the development of this knowledge, as Sasao and Webb (2017) found a relationship between frequency and knowledge. Morphological knowledge is at the intersection of breadth and depth of vocabulary knowledge and needs to be an important learning focus in the beginning and intermediate levels of language learning.

The quality of research on collocation continues to improve, in part helped by the increasing power of computers and the increasing availability of spoken and written corpora of various kinds. A largely neglected area in collocation involves the transparency of collocations, namely how easy is it to get the meaning of the collocation from the meaning of its parts. This is also called compositionality. As argued earlier in this chapter, the transparency of collocations is directly related to how they are comprehended and learned. The categories of core idioms, figuratives, and literals relate to transparency, but within each of these categories there are degrees of transparency. Studies looking at both frequency and transparency could provide data on the size of the learning task for collocations and further refine our understanding of how they might be learned. This could lead to studies of how learners actually cope with them in context.

Testing Control of Aspects of Word Knowledge

There is now a growing number of tests examining learners' knowledge of the written form and the form-meaning connection, and some of these are available in bilingual versions. These include the Vocabulary Size Test (Nation & Beglar, 2007) and the Updated Vocabulary Levels Test (Webb, Sasao, & Ballance, 2017). The Picture Vocabulary Size Test (Anthony & Nation, 2017) uses both spoken and written cues to test knowledge of the most frequent 6,000 words of English and is intended for use with young children. There has been vigorous and very helpful debate and research on the unit of counting, the item format and the role of guessing (including the use of *I don't know*), and the interpretation of such tests. This debate and research will undoubtedly improve the nature and use of such tests and will provide useful guidelines for future tests of different aspects of word knowledge.

Webb and Sasao (2013) have developed carefully constructed tests of word part knowledge. Read's (1995) work on the Word Associates Test combined collocational knowledge and knowledge of associations and encouraged research on the relationship between breadth and depth of vocabulary knowledge (Qian, 1999), with the finding that as vocabulary size (breadth) grows, so does depth of knowledge.

There is a need for tests that look for systematic aspects of word knowledge, such as control of the spelling system, dealing with related senses of words, and the ability to understand figuratives. Some of the measures used with young native speakers, such as the running record (Clay, 2013) where a learner is scored on the ability to read a text aloud, and reading comprehension tests may be adapted to become useful diagnostic tools when working with learners of English as a foreign language.

There is also a need for diagnostic tests of strategies for dealing with aspects of word knowledge that are relevant for foreign language learners. Sasao and Webb (2018) have developed a test to measure skill in guessing from context, which is a strategy for dealing with word meaning. Dictionary use is another such strategy for accessing word meaning, and flash card use is a strategy for learning form-meaning connections. There are form, meaning, and use strategies which can all contribute to knowing words.

There are three important messages to take from this chapter. Firstly, there is more to knowing a word than knowing what it means. Secondly, depth of word knowledge gradually develops in a variety of ways and teaching is only one of the contributors to this knowledge, although not the only one where teachers can have a positive influence. Thirdly, it is possible to monitor the development of many aspects of knowledge through tests and observational procedures, and both teachers and learners should make use of these.

Further Reading

Nation, I. S. P. (2013). Knowing a word. In I. S. P. Nation (Ed.), *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press, Chapter 2, pages 44–91.

This chapter focuses on knowing a word, with an extended discussion of the receptive-productive distinction and the nine aspects involved in knowing a word. There is a substantial list of references for the chapter and these can be updated by referring to the regularly updated vocabulary bibliography that appears on Nation's website, particularly those articles followed by the reference number [2].

Nation, I. S. P. (2013). Finding and learning multiword units. In I. S. P. Nation (Ed.), *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press, Chapter 12, pages 479–513.

This chapter looks critically at research on collocation and multiword units and suggests ways of searching for them in corpora and in classifying them.

Webb, S., & Nation, I. S. P. (2017). Learning burden. In S. Webb & I. S. P. Nation (Eds.), *How vocabulary is learned*. Oxford: Oxford University Press, pages 25–42.

This chapter has a detailed discussion of learning burden and how it can be applied.

Sasao, Y., & Webb, S. (2017). The word part levels test. *Language Teaching Research*, 21(1), 12–30.

This comprehensive study of knowledge of English affixes provides tests of three kinds of affix knowledge plus results that can be used to guide teaching.

Related Topics

The mental lexicon, the relationship between vocabulary knowledge and proficiency, L1 and L2 vocabulary size and growth, how vocabulary fits into theories of L2 learning, incidental vocabulary learning, measuring depth of vocabulary knowledge

References

- Anthony, L., & Nation, I. S. P. (2017). *Picture Vocabulary Size Test* (Version 1.0.0) [Computer Software]. Tokyo, Japan: Waseda University. Retrieved from www.laurenceanthony.net/software/pvst
- Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Language Learning*, 57(1), 35–56.
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge, UK: Cambridge University Press.
- Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Basingstoke, UK: Palgrave Macmillan.
- Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 23, 321–343.
- Clay, M. (2013). *An observation survey of early literacy achievement* (3rd ed.). Portsmouth, NH: Heinemann Publishing.
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System*, 25(3), 301–315.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238.
- Daulton, F. E. (2008). *Japan's built-in lexicon of English-based loanwords*. Clevedon, UK: Multilingual Matters.
- Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge, UK: Cambridge University Press.
- de Groot, A. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning*, 56(3), 463–506.
- Deconinck, J., Boers, F., & Eyckmans, J. (2017). “Does the form of this word fit its meaning?” The effect of learner-generated mappings on L2 word recall. *Language Teaching Research*, 21(1), 31–53.

- Ellis, R. (2005). Principles of instructed language learning. *System*, 33, 209–224.
- Flege, J. E. (1981). The phonological basis of foreign accent: A hypothesis. *TESOL Quarterly*, 15(4), 443–455.
- Grant, L., & Nation, I. S. P. (2006). How many idioms are there in English? *ITL – International Journal of Applied Linguistics*, 151, 1–14.
- Griffin, G. F., & Harley, T. A. (1996). List learning of second language vocabulary. *Applied Psycholinguistics*, 17, 443–460.
- Guiora, A. Z., Beit-Hallami, B., Brannon, R. C. L., Dull, C. Y., & Scovel, T. (1972). The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry*, 13, 421–428.
- Guiora, A. Z., Brannon, R. C. L., & Dull, C. Y. (1972). Empathy and second language learning. *Language Learning*, 22(1), 111–130.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review*, 61(3), 355–382.
- Joe, A. (1998). What effects do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19(3), 357–377.
- Liu, D. (2010). Going beyond patterns: Involving cognitive analysis in the learning of collocations. *TESOL Quarterly*, 44(1), 4–30.
- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, 33(3), 299–320.
- McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11(2), 113–128.
- McLean, S. (2017). Evidence for the adoption of the flemma as an appropriate word counting unit. *Applied Linguistics*, 1–24. <https://doi.org/10.1093/applin/amw050>
- Meara, P. (2006). Emergent properties of multilingual lexicons. *Applied Linguistics*, 27(4), 620–644.
- Miller, G. A., & Fellbaum, C. (1991). Semantic networks in English. *Cognition*, 41, 197–229.
- Mochizuki, M., & Aizawa, K. (2000). An affix acquisition order for EFL learners: An exploratory study. *System*, 28, 291–304.
- Nakata, T. (2011). Computer-assisted second language vocabulary learning in a paired-associate paradigm: A critical investigation of flashcard software. *Computer Assisted Language Learning*, 24(1), 17–38.
- Nation, I. S. P. (2013a). *Learning vocabulary in another language* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Nation, I. S. P. (2013b). *What should every EFL teacher know*. Seoul: Compass Publishing (available free from Paul Nation’s web site).
- Nation, I. S. P. (2016). *Making and using word lists for language learning and teaching*. Amsterdam: John Benjamins.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.
- Nation, I. S. P., & Crabbe, D. (1991). A survival language learning syllabus for foreign travel. *System*, 19(3), 191–201.
- Parent, K. (2012). The most frequent English homonyms. *RELC Journal*, 43(1), 69–81.
- Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition*, 38(1), 97–130.
- Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: *Do Things Fall Apart? Reading in a Foreign Language*, 22(1), 31–55.
- Perfetti, C. A., & Lesgold, A. M. (1979). Coding and comprehension in skilled reading and implications for reading instruction. In L. B. Resnick et al. (Eds.), *Theory and practice of early reading* (pp. 57–84). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), 1–28.
- Pressley, M. (1977). Children’s use of the keyword method to learn simple Spanish vocabulary words. *Journal of Educational Psychology*, 69(5), 465–472.
- Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56(2), 282–307.

- Read, J. (1995). Refining the word associates format as a measure of depth of vocabulary knowledge. *New Zealand Studies in Applied Linguistics*, 1, 1–17.
- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 209–227). Amsterdam: John Benjamins.
- Sasao, Y., & Webb, S. (2017). The word part levels test. *Language Teaching Research*, 21(1), 12–30.
- Sasao, Y., & Webb, S. (2018). The guessing from context test. *ITL – International Journal of Applied Linguistics*, 169(1), 115–141.
- Schmitt, N. (Ed.). (2004). *Formulaic sequences*. Amsterdam: John Benjamins.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, 19, 17–36.
- Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics*, 16, 155–172.
- Sinclair, J. (2004). New evidence, new priorities, new attitudes. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 271–299). Amsterdam: John Benjamins.
- Stevick, E. W. (1978). Toward a practical philosophy of pronunciation: Another view. *TESOL Quarterly*, 12(2), 145–150.
- Wang, M.-T. K., & Nation, P. (2004). Word meaning in academic English: Homography in the Academic Word List. *Applied Linguistics*, 25(3), 291–314.
- Waring, R. (1997). A study of receptive and productive learning from word cards. *Studies in Foreign Languages and Literature (Notre Dame Seishin University, Okayama)*, 21(1), 94–114.
- Webb, S. (2009). The effects of receptive and productive learning of word pairs on vocabulary knowledge. *RELC Journal*, 40(3), 360–376.
- Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford, UK: Oxford University Press.
- Webb, S., & Sasao, Y. (2013). New directions in vocabulary testing. *RELC Journal*, 44(3), 263–277.
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL – International Journal of Applied Linguistics*, 168(1), 34–70.
- Wei, Z. (2015). Does teaching mnemonics for vocabulary learning make a difference? Putting the keyword method and the word part technique to the test. *Language Teaching Research*, 19(1), 43–69.
- West, M. (1953). *A general service list of English words*. London, UK: Longman, Green & Co.
- Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford, UK: Oxford University Press.

Classifying and Identifying Formulaic Language

David Wood

Introduction

Formulaic language (FL) is generally defined as multiword language phenomena which holistically represent a single meaning or function, and are likely mentally stored and used as unanalyzed wholes, as are single words. The phenomenon itself is generally called formulaic language, and items themselves are referred to as formulaic sequences. FL is a rather enigmatic and elusive element of language, in itself it is a relatively recent subject of focus in linguistics and applied linguistics even though key categories and types have been subjects of scrutiny since the early to mid 1900s.

Over a long period of time the multiword units we now call formulaic language were examined more or less in isolation. This is largely due to the fact that researchers were looking at different and relatively discrete categories of multiword units, often working in quite separate areas of linguistics or other fields, including social anthropology and neurology. Over time, the existing research was examined and reinterpreted as a whole body of knowledge, but it was the late 1990s when the term FL came into common use, largely as a result of the work of Wray (e.g., 1999). The term is now standard, and a great deal of important work has been conducted into many aspects of FL and its use.

Critical Issues and Topics

There is a surprising range and scope of types of formulaic language, as seen in detail later. The categories, when examined, show quite a bit of overlap and imprecision, and are subject to interpretation. For example, determining whether a given sequence is a collocation or an idiom is sometimes a challenge. Some items do not fit comfortably in any specific category, or fall into cracks, for example sequences like *and then* or *sooner or later* are really difficult to categorize. Advances in corpus analysis technology and techniques have helped uncover new types of formulaic sequences, but all the same a sort of orthodoxy has been established over time, which can be puzzling. It is unclear, for example, how significant it is to determine a lexical bundle by means of frequency only, as compared to a sequence identified using frequency in combination with other statistical measures. As well, some categories overlap,

and, perhaps most alarmingly, there is no firm consensus that all the categories are similarly processed semantically or psycholinguistically. It is also logical to question whether the categorizations are useful to researchers or teachers. Is it possible that the classifications are really just leftovers from early studies in phraseology, and that they are largely irrelevant to concerns particular to applied research or language teaching?

The identification of formulaic language in spoken or written texts is a challenging enterprise. Formulaic sequences may be identified in corpora by frequency, and the ways in which formulaic sequences are produced also gives us clues as to what multiword combinations might be formulaic. A potentially valuable means of determining formulaicity involves expert or native speaker judgment, especially useful when applied to small or very specific data sets. Although it is often best to try to use a combination of measures, in many cases absolute certainty in identification can be elusive. Even using combinations of corpus frequency and statistical measures of co-occurrence, along with acoustical features and judges, it is common to hedge one's claims about formulaicity. We can hope that new and more reliable or valid means of identifying formulaic language will come along in time.

Classifications

FL has been labeled in many different ways, and nearly 20 years ago it was Wray and Perkins (2000, p. 3) who identified 40 terms. The main categories of FL are collocations, idioms, lexical phrases, lexical bundles, metaphors, proverbs, phrasal verbs, n-grams, conegrams, and compounds. Some of the sequences are characterized mainly by their structural/semantic/syntactic properties, some by their pragmatic utility, and some by their distribution in particular corpora.

Sequences Distinguished by Structural, Semantic, or Syntactic Properties

Collocations

The term collocation has been around for many years, and research has been inspired by the pioneering work of Firth in the 1950s. Collocations likely come into use because of repeated context-dependent use. Such terms as *senior management*, *single parent*, and *plastic surgery* are examples of collocated pairs of words. Collocation basically refers to a syntagmatic relationship among words which co-occur. The relationship may be restricted to relationships which conform to certain syntactic and/or semantic criteria. Collocations can be in a syntactic relationship such as verb + object for example *make a decision*. Two approaches to collocation research have dominated: frequency-based and phraseological (see Granger & Paquot, 2008, for an overview of these). The frequency-based approach is in the tradition of the work of Firth (1951, 1957) and deals with the statistical probabilities of words appearing together, while the phraseological approach, dating back to very early work in Soviet phraseology, is much more concerned with restrictive descriptions of multiword units, and takes quite narrow views of what constitutes a collocation. To add to the complexity of the use of the term collocation, researchers have used it more creatively, sometimes as an umbrella term for multiword units in general. Frequency-based work on collocation was developed by Halliday, Mitchell and Greenbaum, Sinclair, and Kjellmer. These researchers extended and refined the definition to specify that a collocation is a function of the frequency of a word appearing in a certain lexical context as compared to its frequency in language as a whole.

They included syntactic and semantic aspects in descriptions of collocations, and explored the issue of what span of words to consider a collocation. Jones and Sinclair (1974) found that the span of words which is optimal for a collocation is four words to the right or left of a node, or core word. Kjellmer worked on the *Dictionary of English Collocations* (1994) defining a collocation as a continuous and recurring sequence of two or more words which are grammatically well formed. This led to the development of computer-based frequency driven study of collocations.

Unlike the frequency-focused researchers, phraseologists tend to see collocations as multi-word units whose component relations are variable and whose meaning is somewhat transparent (Nesselhauf, 2005). For example, Cowie (1994) placed collocations along a scale from composites, combinations below the sentence level with lexical or syntactic functions (e.g., *red herring*), and formulae, often sentence-length and having pragmatic functions (e.g., *how are you?*). Composites can be fully opaque and/or invariable, as in “pure idioms” (e.g., *kick the bucket*). “Figurative idioms” can have both a literal and figurative meaning (e.g., *to play a part*), and restricted collocations in which at least one element is literal and the other figurative (e.g., *explode a myth*). Cowie gives no restrictions on the number of words or the span of words in a collocation.

Idioms

Idioms are perhaps the archetypal formulaic sequence. Unfortunately, they are as ambiguous as collocations, and share with them a sort of dual personality, with idiom referring both to a specific type of FL, and some researchers using the term more broadly, with definitions encompassing proverbs, slang expressions, and so on. In general, however, the term is used to refer to word combinations which are, in the words of Moon (1998, p. 4), “fixed and semantically opaque or metaphorical”, for example, *kick the bucket* or *spill the beans*.

A key quality of an idiom is its semantic non-compositionality and non-productive form. Wood (1981) noted that the meaning of an idiom is not the sum of the meanings of its component parts, that is, it is not compositional, and its structure must not be transformable, that is, it is non-productive, or frozen. Examples of items which meet these two criteria are *kick the bucket* and *by and large*, which cannot be understood by means of their constituent parts, nor can they be grammatically manipulated. In fact, many types of formulaic sequences display idiomaticity to greater or lesser degrees.

The modern scholar with the most useful definitions and categories of idioms is Moon (1998), who defined idioms as “semi-transparent and opaque metaphorical expressions such as *spill the beans* and *burn one’s candle at both ends*” (p. 5). She differentiated idioms from fixed expressions such as routine expressions, sayings, similes, and so on (Moon, 1998, p. 2). Somewhat later, Grant and Bauer (2004) added the stipulation that an idiom is also non-figurative, meaning that its meaning must not be interpretable from the component words. Applying the criteria of Grant and Bauer, *kill two birds with one stone* is not an idiom because its meaning may be seen as nonliteral, and then interpreted again through examining its pragmatic intent. In contrast, applying the criteria of Grant and Bauer, *by and large* is an idiom, because it is both nonliteral and provides no indication of its figurative meaning.

In sum, there are five basic defining criteria of an idiom (see Skandera, 2004; Wood, 2015):

- 1 An idiom is two or more words in length.
- 2 Semantically opaque (the meaning of the whole is not the sum of the meanings of individual component words) – examples might be *spic and span* and *to and fro*

in which the component words are also opaque, *spic*, *span*, or *fro* are never used outside of these contexts (see Allerton, 1984). Many instances of opaque idioms have historical roots, for example *kick the bucket* (die), relates to the slaughtering of pigs.

- 3 Noncompositionality – the words that make up an idiom cannot be analyzed for meaning or function. This is akin to/linked to semantic opacity.
- 4 Mutual expectancy – this can also be termed *lexicality*, and refers to the fact that the component words of an idiom co-occur in a fixed manner, giving the idiom a unitary form to accompany its holistic meaning or function. It is, in essence, operating as a single lexical item.
- 5 Lexicogrammatical invariability/frozenness/fixedness – the component words in an idiom are fixed and cannot be substituted by synonyms. In fact, some idioms do not even allow syntactic or morphological variation. Examples include *hook line and sinker* or *beat around the bush*; we cannot, for example, pluralize any of the nouns in these sequences, nor, for example, passivize the voice to render another appropriate idiom such as *the bush is beaten around*.

Metaphors

A metaphor is a sequence based on an unconventional reference in which words are used to describe something ordinarily far from its normal scope of denotation, producing a discord between a literal interpretation and a metaphoric interpretation. Metaphors have a common structure: the *vehicle* is the term used in an unusual manner, and the *topic* is the referent of the vehicle. The shared semantic content between the vehicle and topic are the *grounds*. An example might be *time is a healer*, in which *healer* is the vehicle, used in an unconventional sense, and *time* is the topic. In this case the grounds is the view of time acting like a physical remedy or medical practitioner, healing spiritual or emotional injuries and wounds in the same way as physical illnesses and injuries are healed by medicine or nurses. A metaphor can also be a simile, using *is like* or *kind of*. An example would be *life is like a box of chocolates*. The power of metaphor is linked to the semantic distance between vehicle and topic, and the relative explicitness of the vehicle.

Proverbs

Proverbs are generally sentence-length sequences which display an opaque relationship between literal and figurative meanings. Proverbs provide advice and warning (*a stitch in time saves nine*), instruction and explaining (*early to bed and early to rise*), and communicate common experience and observations (*like death and taxes*). They are taken from a store of proverbs shared by a cultural group or community. They are generally brief, direct, have simple syntax, contain elements of metaphor and sometimes dated or archaic structure or words.

Compounds

A compound is a sequence of two words (see ten Hacken, 2004), the second of which usually functions as the *head* or *core* of the compound – for example *desk computer* describes a type of computer and *computer desk* describes a type of desk (see Williams, 1981). The

head represents a type and the nonhead serves to classify the head. There are three forms of compounds:

- 1 **Closed form**, in which the words are written as one, such as *hardcore* or *laptop*
- 2 **Hyphenated form**, in which the lexical items are separated by hyphens, such as *brother-in-law* or *open-handed*
- 3 **Open form**, in which the two words are written separately, such as *grocery store* or *real estate*

Compounds are sometimes written as single words if the unit is strongly lexicalized. The words may be linked by a hyphen and over time become blended orthographically into a single word. Words modified by adjectives, for example, *a metal table*, are different from a compound word, for example, *a coffee table*, in the degree to which the non-headword changes the essential character of the head, or the degree to which the modifier and the noun are inseparable. In the example of *coffee table*, the compound represents a single entity, a particular type of table which is always identified in the same way, whereas the *metal table* is simply a table being described by means of the material from which it is made. The adjective slot in the sequence can be filled by any number of choices.

Phrasal Verbs

Phrasal verbs are distinguished largely by their distinctive structural makeup, lexical verbs combined with a preposition, particle, or both, with often nonliteral meanings, or both literal and figurative interpretations, like idioms. Three structural categories exist:

- 1 **Verb + preposition (prepositional phrasal verbs)**
She quickly *picked up* some Portuguese on her Brazilian vacation.
I happened to *bump into* my former boss on the street.
- 2 **Verb + particle (particle phrasal verbs)**
You can *show that off* at the next party.
I tried not to *cave in* under the stress of the divorce.
- 3 **Verb + particle + preposition (particle-prepositional phrasal verbs)**
He is always *going on about* something or other.
Jane *looks up to* her older brother.

According to Liu (2008, p. 22) there are three fundamental criteria for determining whether an item is a phrasal verb:

- 1 No adverb between the lexical verb and preposition or particle; for example, we cannot say *The kids loaded slowly up on chocolates before we got there*.
- 2 The particle cannot be at the front of a sentence; for example, we cannot say *Up with I am not putting any more outbursts*.
- 3 It cannot exist as only literal in meaning, but needs to have a figurative meaning, as in the preceding examples.

Sequences Distinguished by Pragmatic Utility

Lexical Phrases

Lexical phrases are a pragmatically specialized subset of formulaic sequences first described by Nattinger and DeCarrico (1992). The phrases fall into two structural categories: **strings of specific lexical items**, mostly unitary and grammatically canonical, and **generalized frames**, category symbols and specific lexical items. The phrases display four characteristics: length and grammatical status; canonical or noncanonical shape; variability or fixedness; continuousness or discontinuousness, the latter allowing lexical insertions (Nattinger and DeCarrico, 1992, pp. 37, 38). There are four broad categories of lexical phrases: *polywords*, which function as single words, without variability or lexical insertions (e.g., “for the most part”, “so far so good”); *institutionalized expressions*, sentence-length, invariable, and usually continuous (e.g., “a watched pot never boils”, “nice meeting you”, “long time no see”); *phrasal constraints*, which have variations of lexical and phrase categories, and are generally continuous (e.g., “a ___ ago”, “the ___er the ___er”); and *sentence builders*, which contribute to the construction of full sentences with fillable slots (e.g., “I think that X”, “not only X but Y”) (pp. 38–45). It is clear that this particular taxonomy exhibits considerable overlap with other categories of formulaic language, such as proverbs, idioms, and collocations. The distinguishing feature of lexical phrases is that Nattinger and DeCarrico used pragmatic function as their common characteristic.

Pragmatic Formulas

In pragmatics, *formula* is used to refer to formulaic sequences employed for specific pragmatic purposes (Bardovi-Harlig, 2012). Various terms have been used for the types of sequences which have pragmatic functions in spoken interaction, including conversational routines, pragmatic idioms, speech formulas, routine formulas, situation formulas, and situation-bound utterances. These formulas are pragmalinguistic resources in spoken language and serve as the most socially appropriate means of accomplishing particular pragmatic functions. These include greetings (*how are you, what’s going on, how are things*) and turn-taking (*let me add . . . , I also think . . . , not only that, but . . .*).

Sequences Distinguished by Their Distribution in Corpora

Lexical Bundles

Lexical bundles (see Biber, Johansson, Leech, Conrad, & Finegan, 1999) are formulaic sequences distinguished by the procedures by which they are identified in corpora, and the fact that they are linked purely to functions in discourse, and are not meaning units. Lexical bundles may be considered more a type of multiword unit than strictly formulaic sequences, since there is no indication in the literature that they are stored or retrieved as wholes. They are researched using particular methods which focus exclusively on frequency and function. Research on lexical bundles tends to overwhelmingly focus on academic language, particularly written text.

Lexical bundles can be briefly defined as “combinations of three or more words which are identified in a corpus of natural language by means of corpus analysis software programs” (Wood, 2015, p. 45). Lexical bundles appear in a range of texts in a corpus. They have been

shown to be essential to the construction of academic writing, with particular bundles used more in specific disciplines (Cortes, Jones, & Stoller, 2002).

The pioneer of lexical bundles research is Biber (2006), who discovered that academic disciplines use specific lexical bundles, and created a categorization of functions of bundles – *referential bundles* – which refer to real or abstract entities or to textual content or attributes, for example, “the size of the . . .”, “one of the things”; *stance bundles*, which express attitudes or assessments of certainty, for example, “. . . are likely to be . . .”, “what do you think . . .”; *discourse organizers*, which indicate connections between previous and subsequent discourse, for example, “on the other hand”, “as well as . . .”.

Concgrams

A concgram is a combination of two or more words, but is distinctive in that it is a noncontinuous sequence, with the constituent words separated by others. The COBUILD team at the University of Birmingham in the 1980s were the first to use computer software to search corpora for noncontiguous word sequences. According to Sinclair (2005), it is likely that researchers will similarly uncover new patterns of word sequences, with “intercollocability” and “interparaphrasability”, already entering the picture.

Clearly there is a surprising range and scope of types of formulaic language. The phenomenon is not a unitary construct, and classifications overlap and require considerable interpretation. Looking at a sequence and determining whether it is, for example, a collocation or an idiom, one may experience quite a bit of indecision. Some items may appear to fit with no particular category, for example, sequences like *and then* or *sooner or later* seem to defy labeling. Thanks to corpus analysis technology and techniques we have discovered new types of formulaic sequences.

In any case, formulaic sequences fall into various categories based on their features or usage. The descriptions of categories have evolved over time, and the classifications are somewhat fluid, with plenty of overlap and outliers.

Identification

It is useful to have a definition of formulaic language, and a sense of what the major categories are, because this takes us one step toward being able to handle it in research and in education. However, sooner or later any researcher or educator will come up against an obstacle which is tough to deal with: how can one go about identifying formulaic sequences in texts, spoken or written?

To understand the challenge of identification, examine the first two sentences in this section, and attempt to identify the formulaic sequences. Issues crop up immediately. Several multiword sequences stand out as more or less idiomatic, for example, *one step toward*, *handle it*, *tough to deal with*, *sooner or later*. How confident can one be with these decisions, and what features of the sequences lead us to decide they are formulaic or idiomatic? More importantly, what other elements of the sentences are formulaic but are not readily accessible to our intuitions and perceptions? How can they be uncovered? This question has been a preoccupation in the study of formulaic language. Perhaps one might decide that some sequences are more formulaic than others, but even then, what can guide the decisions? How frequently they are used in a given register? Prosodic features of the production of the sequence? Their frequency in a large corpus?

Fortunately, there are a number of reliable and well-developed means of identifying formulaic sequences. Some are more reliable than others, and they lend themselves to particular purposes and texts.

Frequency and Statistical Measures

It is axiomatic that particular formulaic sequences are generally recurring in a particular register, and a word sequence which sees frequent use is probably formulaic, provided it is also a more or less unitary meaning or function unit. Maybe they will also be mentally stored and retrieved as a single unit.

A distributional or frequency-based approach to identifying formulaic sequences is quite common in research (Durrant & Mathews-Aydinli, 2011). Statistical identification of formulaic language is a very productive method. In this type of identification procedure a set of parameters is established, marked by minimum lengths of sequences, minimum criteria for frequency, usually expressed as occurrences per million words in a corpus. The corpus is scanned for word combinations that fit within the parameters. Frequency cutoffs can range from 10 to 40 occurrences per million words (Biber et al., 1999; Simpson-Vlach & Ellis, 2010). The sequences which are uncovered by this type of corpus-based, statistically driven procedure are often not complete structural units (Cortes, 2004), and the majority of research of this type has uncovered units labeled *lexical bundles* (e.g., Biber et al., 1999) or *multiword constructions* (Liu, 2012; Wood & Appel, 2014). In some cases, researchers have used these parameters as part of a more elaborate process of identification, and have employed the term *formulaic sequences* (e.g., Simpson-Vlach & Ellis, 2010).

Logically, this statistical approach is used mostly with large corpora of hundreds of thousands, if not millions, of words. These corpora are created to be representative of specific registers of language. A purely frequency-based approach has some serious limitations if used with smaller data sets, particularly because minimum frequency cutoffs may be very difficult to set. For example, in a corpus of a million words, a frequency cutoff of 40 occurrences per million words would mean a given sequence would need to occur 40 times in order to be considered formulaic or a lexical bundle. But if a corpus has, say, 100,000 words, the sequence would only need to occur four times. This makes it challenging to identify sequences as formulaic using frequency-based methods alone when working with small corpora. A further limitation of using frequency alone as a criterion for formulaicity is that additional steps are also required to eliminate meaningless combinations of words – some content-specific word combinations can easily pass the frequency test and yet not be formulaic, for example, proper names such as *Ford Motor Company* or verbal tics or nonlexical fillers such as *so, so, so. . .* Furthermore, it would be difficult to rely wholly on frequency for identifying formulaicity transcribed conversations on a range of topics. Many sequences which are formulaic might appear only once or twice in such a diverse and small set of data. A very important drawback of using frequency-based analysis is the fact that frequency gives us absolutely no real indication of the holistic processing, which is often an important concern in some types of research – there is no way to rely on frequency measures to identify formulaic sequences in individual idiolects unless extra means of analysis are used. A good example of this is evident in a study by Schmitt, Grandage, and Adolphs (2004) in which formulaic sequences were identified in a corpus by means of statistical measures. The sequences were then integrated into spoken dictation tasks in which dictated texts were designed to overload the participants' short-term memory capacities. Evidence of holistic

storage of the sequences in the participants' reconstructions of the dictations varied greatly from participant to participant (Schmitt, Grandage, and Adolphs, 2004).

Some word sequences which appear to be formulaic in terms of saliency and unitary meaning or function may not actually appear at particularly high frequency in a given corpus. For example, *in spite of* or *how are you* may not be frequent in any particular corpus or genre, but we would probably agree that they show formulaicity, because they contain words which very commonly occur in this order, and they have a particular unitary meaning or function. The probability that words will co-occur like this can be measured statistically using measures of association such as Mutual Information (MI). MI is a measure of how likely a given set of words are to occur together in a set sequence in comparison to chance. MI does not have a particular statistical significance cutoff, but many researchers have used an MI of 3.0 or higher as an indication of statistical strength of co-occurrence (e.g., Church and Hanks, 1990; Hunston, 2002). A higher MI represents a stronger probability of co-occurrence, and is quite a strong objective measure of formulaicity, if used in combination with other frequency measures. Other similar measures of strength of association of words are used in corpus linguistics. For example, Gries (2008, 2012) uses a measure called the Fisher-Yates exact probability test in examining the strength of the relationships between a given word and a construction in which it occurs. Some studies combine data of various types such as corpus measures of association, eye tracking, and response latency. These types of measures are generally called *psycholinguistic measures*.

When frequency measures are not feasible for analysis of a small or very specific corpus, it is possible to look at a large general corpus, such as the British National Corpus (BNC) or the Corpus of Contemporary American English (COCA). These huge sets of language data can provide a wealth of information about particular word combinations in real-life language use. An example of this is in a study by Wood and Namba (2013) in which they identified useful formulaic sequences to help individual Japanese university students improve their oral presentations. The researchers first identified useful sequences by means of native speaker/proficient speaker intuition, and then turned to the COCA to check their frequencies and statistical strength of co-occurrence. The researchers looked for the sequences in the spoken language subcorpus of the COCA, using a frequency cutoff of at least ten occurrences per million words and with a Mutual Information score (MI) of at least 3.0. In this way, they could be certain that the identified sequences were frequent in speech and that they consisted of words with a high frequency of co-occurrence. A novel means of determining formulaicity is the use of online search engines such as Google. Shei (2008) pointed out that there are really no readily available corpora which are large enough to give full coverage of language use for many types of investigation. Shei presents a strong case for researchers and educators to use the internet as an enormous corpus, readily exploited by means of a search engine such as Google to identify and retrieve word sequences for research and language teaching and learning support. It is a simple matter of Googling a given sequence and examining the resulting hits, which may contain extremely valuable information about its frequency, form, variability, and functions.

Psycholinguistic Measures

Studies in which identification of formulaic language is a focus have used measures of processing speed. These measures may include reaction times (e.g., Conklin & Schmitt, 2012), eye movement (e.g., Underwood, Schmitt, & Galpin, 2004), and electrophysiological (ERP) measures (e.g., Tremblay & Baayen, 2010).

Measures such as eye tracking or response latencies require that participants read, making them unsuitable for research involving children or nonliterate individuals. It is obvious that these psycholinguistic measures are useful in identifying sequences psycholinguistically stored by any one individual, but they are much less helpful in showing how commonly used a formulaic sequence may be in a broader speech community. These measures may reveal formulaic sequences which are rare, unusual, or one-off, used idiosyncratically by a speaker.

Acoustic Analysis

A common criterion for identification of formulaic sequences in speech is *phonological coherence*, a term coined by Peters (1983). Phonological coherence is a characteristic of formulaic production in which a word sequence is uttered fluently, with no hesitations and an unbroken intonation Peters (1983, p. 8). Formulaic sequences exhibit certain prosodic characteristics, such as alignment with pauses and intonation units, resistance to internal dysfluency, no internal hesitations, fast speech rhythm, and stress placement restrictions (see Lin, 2010, 2012). It is important to bear in mind that phonological coherence characterizes formulaic sequences in a given individual's idiolect, and that analysis of this type is restricted by the quantity of data which can be processed by an individual and the technological tools used to record and analyze speech data.

Criteria Checklists and Native-Speaker Intuition

Researchers sometimes discover that frequency, psycholinguistic processing, or acoustic analysis measures are insufficient to identify formulaic sequences in various types of data, especially spoken data. This is often resolved by means of criteria checklists that blend specific features associated with formulaicity.

A proponent of the use of such checklists has been Wray (2002), who reviewed methods of detecting formulaic sequences in many data types. She notes that use of corpus analysis computer software is one possible method of identification, but points out some shortcomings of reliance on frequency in particular:

It seems, on the surface, entirely reasonable to use computer searches to identify common strings of words, and to establish a certain frequency threshold as the criterion for calling a string "formulaic" . . . (however) problems regarding the procedures of frequency counts can be identified. Firstly, corpora are probably unable to capture the true distribution of certain kinds of formulaic sequences. . . . The second serious problem is that the tools used in corpus analysis are no more able to help decide where the boundaries between formulaic sequences fall than native speaker judges are.

(pp. 25, 27, 28)

It is obvious that small data sets composed of spoken discourse present challenges for computer corpus analysis software. For one thing, the discourse or topic-specific speech in such data sets, combined with the small total word count, make it very difficult to rely on frequency alone, since some sequences might occur only once or be used very idiosyncratically. It is also often the case that formulaic sequences blend into surrounding language; many also have large fillable slots, presenting a great challenge for corpus analysis software. Research involving second language learners often produces data with large numbers

of nonstandard or idiosyncratic sequences. In the end, a researcher can turn to one quite daunting measure in identifying formulaicity in language, what Wray terms “the application of common sense” (p. 28).

Native-Speaker Judgment

Fortunately, it is readily possible to apply common sense to the task of identifying formulaic sequences, especially that of second language speakers, by examining language performance and comparing it to native-speaker use of formulaic sequences. This involves the use of native-speaker judgment and a checklist of criteria. Wray (2002, p. 23) points out five challenges inherent in this type of procedure:

- 1 It has to be restricted to smaller data sets.
- 2 Inconsistent judgment may occur due to fatigue or alterations in judgment thresholds over time.
- 3 There may be variation between judges.
- 4 There may not be a single answer as to what to search for.
- 5 Application of intuition in such a way may occur at the expense of knowledge we do not have at the surface level of awareness.

Recall how challenging it was at the beginning of this chapter to isolate formulaic sequences without any guiding criteria. This challenge can be at least partially overcome by use of a **checklist** of specific criteria. The standard procedure for this involves judges studying the criteria which inform a checklist, and examining a corpus to apply the criteria and identify sequences which appear to be formulaic. A high degree of interrater reliability among judges is a good general measure of the strength of a given judgment.

A number of checklists have been used in such research; some checklists developed for specific populations, others more general. Following are descriptions of three such checklists which have been used in various studies: an early checklist elaborated by Coulmas (1979); a checklist applicable to a range of child and adult native or non-native speakers (Wray & Namba, 2003); a checklist used to identify formulaicity in second language acquisition of speech fluency (Wood, 2006, 2009, 2010).

Coulmas, 1979

Coulmas (1979, p. 32) lays out nine specific criteria for formulaicity:

- 1 At least two morphemes long (i.e., two words)
- 2 Coheres phonologically
- 3 Individual elements are not used concurrently in the same form separately or in other environments
- 4 Grammatically advanced compared to other language
- 5 Community-wide formula
- 6 Idiosyncratic chunk
- 7 Repeatedly used in the same form
- 8 Situationally dependent
- 9 May be used inappropriately

Wray and Namba, 2003

Wray and Namba (2003) presented a very flexible and comprehensive checklist, originally used in a study of speech of bilingual children. The checklist is applicable to many types of data and consists of 11 criteria, rated on a Likert Scale of 1 to 5. This is quite refined in that it deals with the issue of gradience or ranges of formulaicity:

- 1 By my judgment, there is something grammatically unusual about this wordstring.
- 2 By my judgment, part or all of the wordstring lacks semantic transparency.
- 3 By my judgment, this wordstring is associated with a specific situation and/or register.
- 4 By my judgment, the wordstring as a whole performs a function in communication or discourse other than, or in addition to, conveying the meaning of the words themselves.
- 5 By my judgment, this precise formulation is the one most commonly used by this speaker/writer when conveying this idea.
- 6 By my judgment, the speaker/writer has accompanied this wordstring with an action, use of punctuation, or phonological pattern that gives it special status as a unit, and/or is repeating something s/he has just heard or read.
- 7 By my judgment, the speaker/writer, or someone else has marked this wordstring grammatically or lexically in a way that gives it special status as a unit.
- 8 By my judgment, based on direct evidence or my intuition, there is a greater than-chance-level probability that the speaker/writer will have encountered this precise formulation before, from other people.
- 9 By my judgment, although this wordstring is novel, it is a clear derivation, deliberate or otherwise, of something that can be demonstrated to be formulaic in its own right.
- 10 By my judgment, this wordstring is formulaic, but it has been unintentionally applied inappropriately.
- 11 By my judgment, this wordstring contains linguistic material that is too sophisticated, or not sophisticated enough, to match the speaker's general grammatical and lexical competence.

Native-Speaker Judgment: Wood, 2010

Wood (2010) published a study examining the possible effect of use of formulaic language on speech fluency in second language learners of English. Identifying formulaic sequences was central to the methods used in the research.

Five criteria were integrated into the checklist. They were used as guides for native-speaker judges, and no one criterion or combination of criteria was required in order for a sequence to be labeled as formulaic.

- 1 **Phonological coherence and reduction.** Formulaic sequences may be produced with phonological coherence (Coulmas, 1979; Wray, 2002), lacking internal pausing and exhibiting a continuous intonation contour. Phonological reduction is also possible, involving phonological fusion, reduction of syllables, deletion of schwa, all of which are characteristic of the most frequent phrases in English (Bybee, 2002).
- 2 **The taxonomy used by Nattinger and DeCarrico (1992).** This taxonomy was described above, and is not required in every case, but as a potential guide to determining formulaicity. If a sequence matched a category in the taxonomy it might be flagged as formulaic.