

www.papyruspub.com
second edition

SECOND LANGUAGE RESEARCH

Methodology and Design

ALISON MACKEY AND SUSAN M. GASS

www.papyruspub.com



Second Language Research

In this second edition of the best-selling *Second Language Research*, Alison Mackey and Susan M. Gass continue to guide students step by step through conducting the second language research process with a clear and comprehensive overview of the core issues in second language research. Supported by a wealth of data examples from published studies, the book examines questions of what is meant by research and what defines good research questions, covering such topics as basic research principles and data collection methods, designing a quantitative research study, and concluding and reporting research findings. Supplementary materials, including an extensive glossary and appendices with helpful materials that students can use in conducting their own studies, serve as useful reference tools, with suggestions on how to get research published re-emphasizing the book's practical how-to approach. The second edition of *Second Language Research* is the ideal resource for understanding the second language research process for advanced undergraduates, graduate students, and professionals in Second Language Acquisition and Applied Linguistics.

Alison Mackey is Professor in the Department of Linguistics at Georgetown University.

Susan M. Gass is University Distinguished Professor in the Department of Linguistics and Germanic, Slavic, Asian and African Languages at Michigan State University.

This page intentionally left blank

Second Language Research

Methodology and Design

Second Edition

**Alison Mackey and
Susan M. Gass**

*Georgetown University and
Michigan State University*

This edition published 2016
by Routledge
711 Third Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

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

© 2016 Taylor & Francis

The right of Alison Mackey and Susan M. Gass to be identified as authors of this work has been asserted by them 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 utilized 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.

First edition published by Lawrence Erlbaum Associates, Inc. 2005

Library of Congress Cataloging in Publication Data

Mackey, Alison, author

Second language research: methodology and design/Alison Mackey and Susan M. Gass, Georgetown University and Michigan State University. — Second Edition.

pages cm

Previous edition: Second language research/Alison Mackey, 1st ed., 2005.

Includes bibliographical references and index.

1. Second language acquisition. 2. Second language acquisition — Research—Methodology. I. Gass, Susan M., author II. Title.

P118.2.M23 2015

418.0072—dc23

2015006149

ISBN: 978-1-138-80855-3 (hbk)

ISBN: 978-1-138-80856-0 (pbk)

ISBN: 978-1-315-75060-6 (ebk)

Typeset in Helvetica Neue and Optima
by Florence Production Ltd, Stoodleigh, Devon, UK

Contents

<i>Preface</i>	xv
1. Introduction to Research	1
1.1 Research Methods	2
1.2 Different Types of Research	3
1.3 What Is a Research Report?	6
1.3.1 <i>Title Page</i>	8
1.3.2 <i>Abstract</i>	8
1.3.3 <i>Introduction</i>	9
1.3.4 <i>Methods Section</i>	11
1.3.4.1 Participants	11
1.3.4.2 Materials	11
1.3.4.3 Procedures	13
1.3.4.4 Analysis	15
1.3.5 <i>Results</i>	16
1.3.6 <i>Discussion/Conclusion</i>	18
1.3.7 <i>Notes</i>	18
1.3.8 <i>References</i>	19
1.3.9 <i>Appendices</i>	19
1.4 Identifying Research Questions	19
1.4.1 <i>Feasibility</i>	22
1.4.2 <i>Research Questions and Hypotheses</i>	23
1.4.3 <i>Replication</i>	26
1.5 Conclusion	27
Points to Remember	27
More to Do and More to Think About . . .	28

2. Issues Related to Data Gathering	30
2.1 Ethical Issues in Research Involving Human Subjects	30
2.1.1 <i>Obtaining Informed Consent from Second Language Learners</i>	31
2.1.1.1 Sufficient Information	32
2.1.1.2 Is Withholding Information Ever Necessary?	35
2.1.1.3 Participant Comprehension in Informed Consent	37
2.1.1.4 Voluntary Participation and Informed Consent	39
2.1.1.5 The Informed Consent Document	43
2.1.2 <i>History of Institutional Review of Human Subjects Research, Compliance, and Problem-Solving</i>	44
2.1.2.1 Purpose of Reviews and IRB Responsibilities	44
2.1.2.2 Why Guidelines to Protect Human Subjects Were Developed	45
2.1.2.3 Development of Research Codes of Ethics	46
2.1.2.4 Preparing a Protocol for the IRB	47
2.2 Conclusion	48
Points to Remember	48
More to Do and More to Think About . . .	49
3. Common Data Collection Measures	52
3.1 Pilot Testing	52
3.2 The Significance of Data Collection Measures	53
3.3 Researching Formal Models of Language	57
3.3.1 <i>Acceptability Judgments</i>	57
3.3.1.1 Materials	59
3.3.1.2 Procedures	61
3.3.2 <i>Elicited Imitation</i>	65
3.3.3 <i>Magnitude Estimation</i>	66
3.3.4 <i>Truth-Value Judgments and Other Interpretation Tasks</i>	68
3.3.5 <i>Sentence Matching</i>	70
3.4 Processing Research	71
3.4.1 <i>Sentence Interpretation</i>	71
3.4.2 <i>Reaction Time</i>	72
3.4.3 <i>Moving Window</i>	73
3.4.4 <i>Eye-tracking</i>	74

3.5	Interaction-Based Research	75
	3.5.1 <i>Picture Description Tasks</i>	76
	3.5.2 <i>Spot the Difference</i>	77
	3.5.3 <i>Jigsaw Tasks</i>	81
	3.5.4 <i>Consensus Tasks</i>	81
	3.5.5 <i>Consciousness-Raising Tasks</i>	84
	3.5.6 <i>Computer-Mediated Research</i>	84
3.6	Strategies and Cognitive Processes	85
	3.6.1 <i>Observations</i>	85
	3.6.2 <i>Introspective Measures</i>	86
	3.6.2.1 <i>Stimulated Recall</i>	87
	3.6.2.2 <i>Think-Alouds or On-Line Tasks</i>	89
	3.6.2.3 <i>Immediate Recalls</i>	94
3.7	Sociolinguistic/Pragmatics-Based Research	95
	3.7.1 <i>Naturalistic Settings</i>	95
	3.7.2 <i>Elicited Narratives</i>	96
	3.7.2.1 <i>Silent Film</i>	97
	3.7.2.2 <i>Film Strips with Minimal Sound</i>	98
	3.7.2.3 <i>Picture Tasks</i>	98
	3.7.3 <i>Discourse Completion Test (DCT)</i>	99
	3.7.4 <i>Role Play</i>	101
	3.7.5 <i>Video Playback for Interpretation</i>	101
3.8	Questionnaires and Surveys	102
3.9	Existing Databases	106
	3.9.1 <i>CHILDES</i>	106
	3.9.2 <i>Other Corpora</i>	106
	3.9.3 <i>IRIS</i>	107
3.10	Conclusion	108
	Points to Remember	108
	More to Do and More to Think About . . .	109
4.	Coding	112
4.1	Preparing Data for Coding	112
	4.1.1 <i>Transcribing Oral Data</i>	113
	4.1.1.1 <i>Transcription Conventions</i>	113
	4.1.1.2 <i>Transcription Machines</i>	116
	4.1.1.3 <i>Technology and Transcription</i>	116
4.2	Data Coding	117
	4.2.1 <i>Scales of Measurement</i>	117
	4.2.2 <i>Coding Nominal Data</i>	118
	4.2.3 <i>Coding Ordinal Data</i>	120
	4.2.4 <i>Coding Interval Data</i>	121

4.3	Coding Systems	122
4.3.1	<i>Common Coding Systems and Categories</i>	123
4.3.1.1	T-Units	123
4.3.1.2	Suppliance in Obligatory Contexts (SOC)	125
4.3.1.3	CHAT	126
4.3.2	<i>Custom-Made Coding Systems</i>	127
4.3.2.1	Question Formation	127
4.3.2.2	Negative Feedback	131
4.3.2.3	Classroom Interaction	131
4.3.2.4	Second Language Writing Instruction	134
4.3.2.5	Task Planning	137
4.3.3	<i>Coding Qualitative Data</i>	137
4.4	Inter-Rater Reliability	139
4.4.1	<i>Calculating Inter-Rater Reliability</i>	140
4.4.1.1	Simple Percentage Agreement	140
4.4.1.2	Cohen's Kappa	140
4.4.1.3	Additional Measures of Reliability	140
4.4.1.4	Good Practice Guidelines for Inter-Rater Reliability	141
4.4.1.5	How Data Are Selected for Inter-Rater Reliability Tests	141
4.4.1.6	When to Carry out Coding Reliability Checks	142
4.5	The Mechanics of Coding	143
4.5.1	<i>How Much to Code?</i>	143
4.5.2	<i>When to Make Coding Decisions?</i>	144
4.6	Software for Coding	145
4.7	Conclusion	147
	Points to Remember	148
	More to Do and More to Think About . . .	148
5.	Research Variables, Validity, and Reliability	150
5.1	Introduction	150
5.2	Hypotheses	150
5.3	Variable Types	152
5.3.1	<i>Independent and Dependent Variables</i>	154
5.3.2	<i>Moderator Variables</i>	155
5.3.3	<i>Intervening Variables</i>	155
5.3.4	<i>Control Variables</i>	156
5.4	Operationalization	158
5.5	Validity	158
5.5.1	<i>Content Validity</i>	159

5.5.2	<i>Face Validity</i>	159	
5.5.3	<i>Construct Validity</i>	159	
5.5.4	<i>Criterion-Related Validity</i>	160	
5.5.5	<i>Predictive Validity</i>	160	
5.5.6	<i>Internal Validity</i>	160	
	5.5.6.1 Participant Characteristics	161	
	5.5.6.2 Participant Mortality	163	
	5.5.6.3 Participant Inattention and Attitude	165	
	5.5.6.4 Participant Maturation	166	
	5.5.6.5 Data Collection: Location and Collector	167	
	5.5.6.6 Instrumentation and Test Effect	167	
5.5.7	<i>External Validity</i>	171	
	5.5.7.1 Sampling	172	
	5.5.7.2 Representativeness and Generalizability	175	
	5.5.7.3 Collecting Biodata Information	177	
5.6	Reliability		180
5.6.1	<i>Rater Reliability</i>	181	
	5.6.1.1 Inter-Rater and Intra-Rater Reliability	181	
5.6.2	<i>Instrument Reliability</i>	182	
	5.6.2.1 Test-Retest	182	
	5.6.2.2 Equivalence of Forms	182	
	5.6.2.3 Internal Consistency	182	
5.7	Conclusion		183
	Points to Remember		183
	More to Do and More to Think About . . .		184
6.	Designing a Quantitative Study		188
6.1	Introduction		188
6.2	Research Materials		191
6.3	Intact Classes		194
6.4	Counterbalancing		196
6.5	Research Design Types		198
	6.5.1 <i>Correlational (Associational) Research</i>	198	
	6.5.2 <i>Experimental and Quasi-Experimental Research</i>	199	
	6.5.2.1 Comparison Group Design	199	
	6.5.2.2 Control Group Design	201	
	6.5.3 <i>Measuring the Effect of Treatment</i>	201	
	6.5.3.1 Pre-Test/Post-Test Design	202	
	6.5.3.2 Post-Test Only Design	203	
	6.5.4 <i>Repeated Measures Design</i>	204	

6.5.5	<i>Factorial Design</i>	205
6.5.6	<i>Time-Series Design</i>	207
6.5.7	<i>One-Shot Designs</i>	210
6.5.8	<i>Meta-Analyses</i>	211
6.6	Conclusion	212
	Points to Remember	212
	More to Do and More to Think About . . .	213
7.	Qualitative Research	215
7.1	Defining Qualitative Research	215
7.2	Gathering Qualitative Data	217
7.2.1	<i>Ethnographies</i>	219
7.2.1.1	Advantages	221
7.2.1.2	Caveats	222
7.2.2	<i>Case Studies</i>	222
7.2.2.1	Advantages	224
7.2.2.2	Caveats	224
7.2.3	<i>Interviews</i>	225
7.2.3.1	Advantages	225
7.2.3.2	Caveats	226
7.2.4	<i>Observations</i>	227
7.2.4.1	Advantages	227
7.2.4.2	Caveats	228
7.2.5	<i>Diaries/Journals</i>	228
7.2.5.1	Advantages	229
7.2.5.2	Caveats	230
7.3	Analyzing Qualitative Data	230
7.3.1	<i>Credibility, Transferability, Confirmability, and Dependability</i>	231
7.3.2	<i>Triangulation</i>	233
7.3.3	<i>The Role of Quantification in Qualitative Research</i>	234
7.4	Conclusion	234
	Points to Remember	235
	More to Do and More to Think About . . .	236
8.	Classroom Research	238
8.1	Classroom Research Contexts	238
8.2	Common Techniques for Data Collection in Classroom Research	239
8.2.1	<i>Observations</i>	239

8.2.1.1	Conducting Classroom Observations	240
8.2.1.2	Observation Procedures and Coding Schemes	243
8.3	Introspective Methods in Classroom Research	253
8.3.1	<i>Uptake Sheets</i>	254
8.3.2	<i>Stimulated Recall</i>	254
8.3.3	<i>Diary Research in Classroom Contexts</i>	256
8.4	Practical Considerations in Classroom Research	258
8.4.1	<i>Logistical Issues to Consider When Carrying out Classroom Research</i>	258
8.4.2	<i>Problematics</i>	261
8.4.2.1	Informed Consent	261
8.4.2.2	Debriefing Participants and Facilitators	262
8.4.2.3	Ensuring Confidentiality and Minimizing Disruption	262
8.4.2.4	Data Segmentation and Coding	263
8.4.2.5	Considering the Instructional Setting	263
8.4.2.6	Summary of Problematics	264
8.5	Purposes and Types of Research Conducted in Classroom Settings	265
8.5.1	<i>The Relationship between Instruction and Learning in Second Language Classrooms</i>	265
8.5.2	<i>Action Research</i>	268
8.5.2.1	Definitions	268
8.5.2.2	Theory and Background to Action Research	269
8.5.2.3	Action Research in Practice	269
8.5.3	<i>Aptitude-Treatment Interaction</i>	272
8.6	Conclusion	272
	Points to Remember	273
	More to Do and More to Think About . . .	274
9.	Mixed Methods	275
9.1	Introduction	275
9.2	What Are Mixed Methods?	277
9.3	Why Use Mixed Methods?	278
9.4	Types of Mixed-Methods Studies	280
9.5	Conducting a Mixed-Methods Study	283
9.5.1	<i>Deciding Whether or Not to Conduct a Mixed-Methods Study</i>	284
9.5.2	<i>Planning Phase</i>	284
9.6	Analyzing Data from a Mixed-Methods Study	287

9.7	Software for Analysis	288
9.8	Conclusion	288
	Points to Remember	288
	More to Do and More to Think About . . .	289
10.	Analyzing Quantitative Data	292
10.1	Introduction	292
10.2	Descriptive Statistics	292
	10.2.1 Measures of Frequency	293
	10.2.2 Measures of Central Tendency	297
	10.2.2.1 Mode	297
	10.2.2.2 Median	297
	10.2.2.3 Mean	297
	10.2.2.4 Outliers	299
	10.2.3 Measures of Dispersion	301
10.3	Normal Distribution	304
10.4	Standard Scores	306
10.5	Probability	307
10.6	Inferential Statistics	311
	10.6.1 Prerequisites	312
	10.6.1.1 Standard Error of the Mean	312
	10.6.1.2 Standard Error of the Difference between Sample Means	312
	10.6.1.3 Degrees of Freedom	312
	10.6.1.4 Critical Values	313
	10.6.1.5 One-Tailed versus Two-Tailed Hypotheses	313
	10.6.2 Parametric versus Nonparametric Statistics	314
	10.6.3 Parametric Statistics	315
	10.6.3.1 T-Tests	315
	10.6.3.2 Analysis of Variance (ANOVA)	317
	10.6.3.3 Two-Way ANOVA	318
	10.6.3.4 Analysis of Covariance (ANCOVA)	318
	10.6.3.5 Multivariate Analysis of Variance (MANOVA)	320
	10.6.3.6 Repeated Measures ANOVA	320
	10.6.4 Nonparametric tests	321
	10.6.4.1 Chi Square (χ^2)	321
	10.6.4.2 Mann-Whitney U/Wilcoxon Rank Sums	322
	10.6.4.3 Kruskal-Wallis/Friedman	323
10.7	Statistical Tables	323
10.8	Strength of Association	325

10.9	Eta ² and Omega ²	325
10.10	Correlation	325
	10.10.1 <i>Pearson Product-Moment Correlation</i>	327
	10.10.1.1 <i>Linear Regression</i>	330
	10.10.1.2 <i>Multiple Regression</i>	331
	10.10.2 <i>Spearman Rho/Kendall Tau</i>	332
	10.10.3 <i>Factor Analysis</i>	332
10.11	New Approaches	332
	10.11.1 <i>Power Analyses</i>	334
	10.11.2 <i>Effect Size</i>	334
	10.11.3 <i>Confidence Intervals</i>	335
10.12	Meta-Analyses	336
10.13	Statistical Packages	337
	10.13.1 <i>SPSS</i>	
	10.13.2 <i>VARBRUL</i>	338
10.14	Conclusion	338
	Points to Remember	338
	More to Do and More to Think About . . .	339
11.	Concluding and Reporting Research	344
11.1	The Importance of Reporting Research	344
11.2	The Final Stages in Reporting Quantitative Research	345
	11.2.1 <i>The Discussion</i>	347
	11.2.2 <i>Limitations, Future Research, and Conclusion Sections</i>	350
11.3	The Final Stages in Reporting Qualitative Research	353
11.4	Reporting Mixed-Methods Research	356
11.5	Checklist for Completing Reports of Research	357
	11.5.1 <i>The Research Problem and Questions</i>	357
	11.5.2 <i>The Research Hypotheses</i>	358
	11.5.3 <i>The Audience</i>	359
	11.5.4 <i>The Abstract</i>	359
	11.5.5 <i>The Literature Review</i>	360
	11.5.6 <i>The Design of the Study</i>	360
	11.5.7 <i>Logistics</i>	361
	11.5.8 <i>Participants</i>	361
	11.5.9 <i>Data Gathering</i>	362
	11.5.10 <i>Data Analysis</i>	362
	11.5.10.1 <i>Transcription</i>	362
	11.5.10.2 <i>Coding Systems</i>	362
	11.5.10.3 <i>Inter-Rater Reliability</i>	363

11.5.10.4 Data Organization	363
11.5.10.5 Statistics	363
11.5.10.6 Presentation	364
11.5.11 Conclusions	364
11.5.12 References	364
11.5.13 Footnotes, Endnotes, Figures, and Tables	366
11.5.13.1 Footnotes and Endnotes	366
11.5.13.2 Figures	366
11.5.13.3 Tables	367
11.5.14 Author's Note/Acknowledgments	367
11.5.15 Post-Research Concerns	367
11.5.16 Final Touches and Formatting	368
11.6 Conclusion	369
Points to Remember	370
More to Do and More to Think About . . .	370
Appendix A: Sample Short Form Written Consent Document for Subjects Who Do Not Speak English	372
Appendix B: Consent to Participate in Research	374
Appendix C: Sample Consent Form for a Study in a Foreign Language Context	376
Appendix D: Sample Consent Form for a Classroom Study	378
Appendix E: Sample Transcription Conventions: "Jeffersonian" Transcription Conventions	381
Appendix F: Sample Transcription Conventions for the L2 Classroom	384
Appendix G: Commonly Used Formulae	387
Glossary	389
References	405
Subject Index	425

Preface

This book is designed to be used as a text for introductory courses on research methodology and design, as well as for other courses in second language studies, where there is an emphasis on research. It can also be used as a general resource book by people carrying out second language research.

The first edition was published a decade ago. This new edition provides updates to reflect changes and developments in the field, together with some internal reorganization, and a new chapter on mixed-methods research, reflecting the current state of the art in the field. We have also included some new features to improve its pedagogically friendly nature, including boxed “time to think” and “time to do” suggestions throughout each chapter, to supplement the exercises at the end of each chapter. Throughout the revisions, we retained our primary goal of providing an introductory text for novice researchers. We explain key concepts and provide concrete examples wherever possible for those with little or no research experience. Exercises are provided throughout the text to allow students to think about the concepts introduced and to get hands-on practice at doing the various activities involved in research. We assume that our readers will have some background in the general topic of second language learning. The discussion and data-based questions throughout each chapter and the activities at the end of each chapter are aimed to promote better understanding of the concepts as readers work through the book. We also include a detailed and updated glossary to aid researchers who prefer to use the book more as a resource than a text.

We take a broad and inclusive view of “second language” research. For this reason, our examples reflect concepts from a variety of perspectives in the second language field. The book is designed to address issues important for research in both second and foreign language settings, for child as well as adult

second language learning, for research on bilingual and multilingual language learning, as well as the acquisition of third and subsequent languages. We have attempted to cast a similarly wide net in our coverage of topics; for example, we include research design issues that range from the use of highly experimental data elicitation tools, to qualitative concerns, as well as teacher-initiated research in classrooms. We also include topics of recent interest in the field, such as dealing with university, institutional, and school review boards that grant permission for data gathering from human subjects, including recent concerns about the replication of research. Although our goal is to acquaint readers with the basic issues, problems, and solutions involved in conducting second language research, we believe that some of the content of the book is also relevant to a wider applied linguistics context. In other words, some issues of design are common to many areas of applied linguistics research, even though the examples may not always be. We also recognize that some chapters might not be relevant to all courses on second language research. The book is designed so that chapters or parts of chapters can be skipped. The most obvious example is Chapter 2, which deals with obtaining consent and obtaining permission from institutional review boards. Not all countries or research contexts require stringent procedures, and sections of this chapter, while important for research in some parts of the world, are less relevant for research in other countries.

Although the book focuses specifically on issues of research design and methodology, we have included one chapter that focuses on introductory statistics. Because the field of statistics is so broad and has its own specialized texts and courses, we provide only a simple overview of some of the basic concepts in this area. For those who intend to conduct detailed statistical analyses, we recommend coursework, expert consultations, and other comparable means of learning about advanced statistics, including statistics textbooks. We do not include specific recommendations about particular statistics texts because the selection of the text depends on the focus of the research problem. Second language research can focus on educational or pedagogical practice or on theory building; it can address issues from a variety of perspectives, including psychology, sociology, linguistics, and bilingualism. We suggest that users of this book consult one of the many appropriate statistics books available.

It is always difficult to decide on the order in which to present information. One researcher's ordering of material and chapters might not coincide with the preferences of another researcher or reader. We have placed information on data gathering at the beginning of the book due to the fact that our experience in teaching research methods courses over the years has led us to believe that researchers need to think about where data come from at the outset of a project, and to think about how data are gathered before becoming immersed in some

of the more technical issues of design. In this book, then, issues of data gathering serve as an anchor for later chapters. Of course, when using the book as a text, we imagine that instructors will adapt the book and reorder chapters to match their particular syllabus and preference for presentation. For this reason, we have made sure, where possible, that each chapter can work as a stand-alone introduction to the area it covers.

We are grateful to many individuals for their support in this project that ended up, like most projects of this sort, having a longer history than we had originally anticipated. For both editions, we first thank the many students we have had in different classes over the years who have not hesitated to provide feedback on our various syllabi and our sequencing of materials, as well as the designs of our own research. The following individuals provided various kinds of invaluable assistance and feedback on the first edition, and we recognize them here: Rebekha Abbuhl, Rebecca Adams, Zoltán Dörnyei, Rod Ellis, Seon Jeon, Kendall King, Patsy Lightbown, Kimberly McDonough, Kara Morgan-Short, Jenefer Philp, Charlene Polio, Rebecca Sachs, Ildikó Svetics, Ian Thornton, and Harriet Wood. In this second edition, we were fortunate, once again, to have the invaluable input and help of our research assistants: Lara Bryfonski and Alex Marsters at Georgetown University, and Lorena Valmori at Michigan State University. Ina Choi of Michigan State University helped with NVivo examples. Luke Plonsky also helped with many parts of this revised edition. His input was essential in helping us to see how to present material better and even to help us better understand particular concepts. After many years of using this book in our own courses and hearing from students (and faculty) around the world, we have made adjustments to this edition based on their comments. Four external reviews of the first edition were commissioned, with extensive comments on the book, some dealing with ways to update the manuscript, some with ways to make things clearer, and some with ordering of material. You know who you are (we do not), and we thank you sincerely for your input. You will undoubtedly see your many helpful suggestions reflected in this edition.

Finally, Leah Babb-Rosenfeld and Elysse Preposi of Routledge have been unwavering in their support of this book and enormously patient awaiting this revised edition. We are grateful for the general support and encouragement we consistently receive from everyone at Routledge.

Alison Mackey
Georgetown, Washington, DC
Susan Gass
Williamston, Michigan

This page intentionally left blank

CHAPTER ONE

Introduction to Research

How do we identify good research questions? How do we answer them? What do we mean by research? These sorts of questions are not always easy to answer, but we anticipate that by the end of this book, you will be in a better position to think about them. The book is practical in nature, aimed at those who are involved in second language studies, second/foreign language instruction and researcher training, and in it, we aim to demystify the research process.

Oxford Dictionaries defines *research* as “*the systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions*” (www.oxforddictionaries.com/us/definition/american_english/research?searchDictCode=all, retrieved January 17, 2015). Thus, in its most basic and simplest form, research is a way of finding out answers to questions.

We are all involved in research every day. Let’s think about what probably occurs as part of many of our lives—being stuck in a traffic jam. As we find ourselves not moving on a freeway, we ask why this has happened and come up with a hypothesis (because there is an accident ahead, for example, or because it is 5:00 p.m. on a Friday afternoon). We then seek verification of our hypothesis by waiting patiently (or impatiently) until the traffic starts moving again. If we see an accident, hear a radio report, look at live traffic on our GPS, or if we see the flashing lights of an emergency vehicle, we might be able to confirm or reject our hypothesis. In the absence of an accident, we might conclude that it must be typical rush hour traffic. In other words, every day we ask questions, come up with hypotheses, and seek confirmation of those hypotheses. Research, then, is not something to be wary of; it is not something that is done only in laboratories or controlled conditions. It is something that we do on a small or large scale every day. We can pose and (hopefully) answer questions of deep theoretical significance and we can also pose questions that assist us in our

daily lives. As teachers, research helps us think through problems that intrigue us with regard to our students and, as a consequence, helps us be more effective.

TIME TO THINK ...

Come up with two to three questions about language learning and/or language teaching that you have been puzzled about. Keep these in mind (or develop new ideas) as you continue to use this book.

In this chapter, we discuss the process of generating research questions and formulating hypotheses, and we outline what readers can expect from a typical research report. We conclude by discussing issues of feasibility and the importance of confirming findings, a process known as replication, in second language research.

1.1 RESEARCH METHODS

Research is not monolithic. That is, there is no single way to go about doing research. The way we approach our understanding of language learning will guide us in how we go about collecting information (data) to answer our questions. For example, there are numerous ways to think about second language learning (for example, there are a number of textbooks that cover the wide field of second language acquisition, such as Gass with Behney and Plonsky, 2013). Some researchers conceptualize language as consisting of a set of linguistic abstractions (mental representations). Learning a second language, then, involves learning a new set of mental representations with the focus being on language forms. These are often called formal approaches. Others view language as a sociocultural phenomenon with language learning based not only on internal linguistic factors, as in formal approaches, but on how language forms interact with external factors (contexts for learning). Still others are concerned with how second languages are processed. Are the mechanisms used when processing a first language the same as those used when processing a second? Each of these approaches (and others) will require different data sets to answer the questions posed. For example, in approaches where the social setting is important, context must be provided. In approaches where only grammatical forms are of concern, context is not a consideration. These general orientations influence the methodologies that are used. In the following section, we outline

two broad approaches to research that have received attention in the second language literature: quantitative and qualitative, and we also discuss mixed-methods studies, which use both. We believe that there is no perfect approach; there are strengths and weaknesses to all. What is important to recognize is that choices (and typically trade-offs) have to be made when conducting research. What is equally important to recognize is that whatever research method we choose, we need to justify using that particular methodology in light of our research question. In other words, the process of conducting research involves theoretical conceptualizations as well as practical considerations. These interrelated notions guide decision-making at all stages of the research process, as will become apparent throughout this book.

TIME TO THINK ...

If you have taken a course on second language acquisition (SLA), what theoretical approach do you align with to the greatest extent?

If you have taught language, what learning approach is most helpful to you in understanding how your students are learning?

1.2 DIFFERENT TYPES OF RESEARCH

As mentioned above, there are many approaches to dealing with research. Two of the most common are known as quantitative and qualitative, although this distinction is somewhat simplistic as the relationship is best thought of as a continuum of research types, and mixed-methods research involves elements of both. Quantitative research generally starts with an experimental design in which a specific hypothesis precedes the quantification of data with follow-up numerical analyses (e.g., a study comparing student test results before and after an instructional treatment). Qualitative studies, on the other hand, generally are not set up as experiments; the data cannot be easily quantified (e.g., a diary study in which a student keeps track of his or her attitudes during a year-long Japanese language course), and the analysis is interpretive rather than statistical. As mentioned above, this is rather a simplistic view because one can imagine a number of variations on this theme. In general, though, quantitative and qualitative research can be characterized as in Table 1.1 (see Reichardt & Cook, 1979).

In this book, we attempt to be as inclusive as possible and cover a variety of research orientations. In particular, we show in Chapter 9 how the types

TABLE 1.1 Characteristics of quantitative and qualitative research

<i>Quantitative Research</i>	<i>Qualitative Research</i>
• obtrusive, involving controlled measurement	• naturalistic and controlled observation
• objective and removed from the data	• subjective
• verification-oriented, confirmatory	• discovery-oriented
• outcome-oriented	• process-oriented
• reliable, involving “hard” and replicable data	• “soft” data
• generalizable	• ungeneralizable, single case studies
• assuming a stable reality	• assuming a dynamic reality
	• close to the data

represented in Table 1.1 are combined; in other words, these represent “pure” forms of these two research types.

Grotjahn (1987) points out that there are many parameters that can be used to distinguish between research types, including the type of data (quantitative or qualitative), the method of analysis (interpretative or statistical), and the manner of data collection (experimental or non-experimental [naturalistic]). He outlines six “mixed” forms (see Table 1.2).

TABLE 1.2 Six mixed forms of research

<i>Type of Research</i>	<i>Form of Data</i>		<i>Method of Analysis</i>		<i>Manner of Data Collection</i>	
	<i>Quantitative</i>	<i>Qualitative</i>	<i>Statistical</i>	<i>Interpretative</i>	<i>Experimental/ Quasi-Experimental</i>	<i>Non-Experimental</i>
Experimental-qualitative-interpretative		✓		✓	✓	
Experimental-qualitative-statistical		✓	✓		✓	
Experimental-quantitative-interpretative	✓			✓	✓	
Exploratory-qualitative-statistical		✓	✓			✓
Exploratory-quantitative-statistical	✓		✓			✓
Exploratory-quantitative-interpretative	✓			✓		✓

TIME TO DO ...

Part 1

Read the following abstract from Philp (2003, p. 99).

Interaction has been argued to promote noticing of L2 form in a context crucial to learning—when there is a mismatch between the input and the learner’s interlanguage (IL) grammar (Gass & Varonis, 1994; Long, 1996; Pica, 1994). This paper investigates the extent to which learners may notice native speakers’ reformulations of their IL grammar in the context of dyadic interaction. Thirty-three adult ESL learners worked on oral communication tasks in NS-NNS pairs. During each of the five sessions of dyadic task-based interaction, learners received recasts of their nontargetlike question forms. Accurate immediate recall of recasts was taken as evidence of noticing of recasts by learners. Results indicate that learners noticed over 60–70 percent of recasts. However, accurate recall was constrained by the level of the learner and by the length and number of changes in the recast. The effect of these variables on noticing is discussed in terms of processing biases. It is suggested that attentional resources and processing biases of the learner may modulate the extent to which learners “notice the gap” between their nontargetlike utterances and recasts.

Does this abstract seem more part of a quantitative or a qualitative study? In thinking about this, consider the following:

- Does the study have quantitative data?
- How are data analyzed?
- Were data collected experimentally?

Part 2

Read the following abstract from Willett (1995, p. 473).

This ethnographic report “thickly describes” (Geertz, 1973) the participation of ESL children in the daily classroom events of a mainstream first-grade classroom. Data for this paper come from a year-long study of one classroom in an international school on a college campus in the U.S. Using a language socialization and micropolitical orientation, the report describes how, through socially significant interactional routines, the children and other members of the classroom jointly constructed the ESL children’s identities, social relations, and ideologies as well as their communicative competence

in that setting. The sociocultural ecology of the community, school, and classroom shaped the kinds of microinteractions that occurred and thus the nature of their language learning over the course of the year.

Does this abstract seem to describe a quantitative or a qualitative study? Consider the following:

- Does the study use naturalistic data?
- If so, of what sort?
- Does it provide an interpretative or a statistical analysis?
- Is there an experimental design?

Part 3

Look at a recent issue of one of the journals listed below. Consider the abstracts for two to three of the articles and determine whether they seem to be more quantitative or qualitative.

- *Studies in Second Language Acquisition*
- *TESOL Quarterly*
- *Language Learning*

What led you to that conclusion?

1.3 WHAT IS A RESEARCH REPORT?

In this section, we provide a guide for readers as to what to expect in a typical article in the field of SLA, focusing primarily on quantitatively oriented research articles. Unlike quantitative research reports where there is a relatively standard format for reporting, qualitative research articles are more wide-ranging in terms of organization (see Chapter 7, where we discuss qualitative research). Our goal is to give an idea of what to expect in a research report. To that end, below we present a basic skeleton of a research paper. Chapter 11 provides detailed information for researchers concerning the *writing* and *reporting* of their own research based on all of the areas covered in this book.

We now consider in more detail what might be included in each of these parts.

Typical research paper format

TITLE PAGE

ABSTRACT

BODY

I. Introduction

- A. Statement of topic area
- B. Statement of general issues
- C. General goal of paper
- D. Literature review
 - 1. Historical overview
 - 2. Major contributions to this research area
 - 3. Statement of purpose including identification of gaps
 - 4. Hypotheses

II. Method

- A. Participants
 - 1. How many?
 - 2. Characteristics (male/female, proficiency level, native language, etc.)
- B. Materials
 - 1. What instruments?
 - 2. What sort of test? What sort of task?
- C. Procedures
 - 1. How is the treatment to be administered?
 - 2. How/when is the testing to be done?
- D. Analysis
 - How will the results be analyzed?

III. Results

Charts, tables, and/or figures accompanied by verbal descriptions

IV. Discussion/Conclusion (often two separate sections)

Common features:

- Restatement of the main idea of the study
- Summary of the findings
- Interpretation of the findings in light of the research questions
- Proposed explanation of the findings, usually including information about any findings that were contrary to expectations
- Limitations of the study
- Suggestions for future research

NOTES

REFERENCES

APPENDICES

TIME TO DO ...

There are many journals where research in SLA is published. Conduct a library or online search and come up with a list of five journals focusing on some area (general or specific) of second language research.

1.3.1 Title Page

- Name of author(s)
- Title of paper
- Contact information

1.3.2 Abstract

The abstract presents a summary of the topic of the paper and the major findings of the research. Abstracts are very often printed through abstracting services and are generally the primary source in finding out about a paper. They are usually 100–150 words in length, although there is variation depending on where the article is published. In the box below is an example of an abstract (Williams, 1999, p. 583).

Abstract

Recent studies have suggested that the incorporation of some attention to form into meaning-centered instruction can lead to improved performance in processing input and increased accuracy in production. Most have examined attention to form delivered by instructors or instructional materials. This study examines the production of 8 classroom learners at 4 levels of proficiency to determine the extent to which learners can and do spontaneously attend to form in their interaction with other learners. Results suggest that the degree and type of learner-generated attention to form is related to proficiency level and the nature of the activity in which the learners are engaged. They also indicate that learners overwhelmingly choose to focus on lexical rather than grammatical issues. (118 words)

In this short abstract, two sentences are devoted to past research, with the third sentence informing the reader what this study is about and how it fills a gap in the literature. The final two sentences provide information about what the reader can expect from the results.

TIME TO DO ...

Find abstracts from two different articles in two different journals.
Analyze each in the way that we did above.

1.3.3 Introduction

The introduction sets the scene and provides the reader with background material (statement of topic area and general issues), as well as an outline of the purpose of the research. This is generally followed by a literature review. Some possibilities for literature reviews include the following:

- Historical overview

Example:

In earlier views of the relationship between x and y . . .

- Major players in this research area, including questions, past findings, and controversies

Example:

In 1994, Smith claimed that the relationship between x and y was an important one and went on to show that . . .

However, in a more recent paper, Jones (1995) argued that this relationship could not be valid because . . .

- General goal of the paper

Example:

In this paper, I will argue that Jones's interpretation of Smith's data is incorrect and that when one looks at variable z in the context of x and y , the relationship is indeed valid. I will present data that support Ellis's original interpretation of abc .

- Research questions/hypotheses

In Williams' (1999) article abstracted above, the following research questions are provided after the introduction (p. 591):

Example:

1. Do learners in learner-centered, communicative classrooms spontaneously attend to form?
2. Is proficiency level related to the extent to which they do so?
3. How do learners draw attention to form?
4. When do learners draw attention to form, that is, during what types of activities?
5. What kinds of forms do they attend to?

As can be seen, these questions build on one another. They are not, however, formulated as predictions. Below are specific hypotheses from a different study (Gass & Alvarez-Torres, 2005):

Example:

1. Given that interaction is said to be an attention-drawing device, we predict that the three experimental groups with interaction will perform better than the group with no interaction.
2. Because input and interaction serve different important functions, when there is a combination of conditions (input followed by interaction and interaction followed by input), performance will be better than when only one type of presentation is available.

3. Given Gass's (1997) assumption that interaction serves as a priming device that "readies" learners to utilize follow-up input, the best performance will take place in the group with interaction followed by input.

The amount of detail needed in a literature review will depend on the purpose of the report. For example, a Ph.D. dissertation will generally be as exhaustive as possible. On the other hand, the literature review for a journal article or for a chapter in a book will only address previous research that directly relates to the specific purpose of the research report and might only be about 5–10 pages.

1.3.4 Methods Section

In the methods section, the reader can expect to be informed about all aspects of the study. One reason for this is the later possibility of replication (see section 1.4.3). Another reason is that in order for a reader to come to an informed opinion about the research, he or she needs to have as much detail as possible about what was done.

1.3.4.1 Participants

This section includes information about the participants¹ in a study. For example, how many participants were there? What are their characteristics (e.g., male/female, native language, age, proficiency level, length of residence, amount and type of instruction, handedness)? The characteristics that researchers describe will depend, in part, on the experiment itself. For example, handedness was listed above as a possible characteristic. This would probably be relevant in a study that required participants to press a button on a computer as a response to some stimulus. Most such studies are set up for right-handed individuals, so it might be important to know if the particular setup favored those individuals.

1.3.4.2 Materials

The materials used to conduct the study are usually presented in detail and where there is insufficient space, a new trend in our field, supported by all the major journals, is to encourage the upload of the original materials to the free, online, up- and downloadable, searchable database IRIS (www.iris-database.org). Below is an example of a materials section from an article on deriving meaning from

written context by Dutch children (grades 2, 4, and 6) in their L1 (Fukkink, Blok, & de Glopper, 2001, p. 481).

Target words were selected from a primary-school dictionary (Verburg & Huijgen, 1994), to warrant that relevant concepts would be selected, representative of the words young readers encounter during reading. An initial sample of words with a frequency below 10 per million (Celex, Centre for Lexical Information, 1990) was selected from this dictionary to ensure that no words were used that students were already familiar with. Three judges evaluated the concreteness of the target words, defined as a dichotomy, and words were excluded if the judges did not arrive at a unanimous agreement. A final sample of 12 words was selected, evenly divided into concrete and abstract words. The average word frequency of the words in the sample is 4.4 per million (ranging from 1 to 10 per million). Only morphologically nontransparent words were included, to promote deriving word meaning from (external) context.

Short texts of approximately a hundred words were constructed for each target word. The difficulty level of each text was adjusted to an appropriate level for average readers at the end of grade 2 on the basis of a reading difficulty index (Staphorsius & Verhelst, 1997). The narrative texts contained no explicit clues (e.g., synonyms, antonyms, or description clues). Target words were not placed in the first sentences of the text.

A version of the twelve texts was presented to three adults with target words deleted. They were instructed to fill each cloze with an answer that was as specific as possible and fitted the context. Only four out of the 36 answers, each concerning a different target word, did not match the concept of the deleted word. The other answers, however, were identical or synonymous with the deleted target word (58 percent) or closely related hypernyms (31 percent) (“to break” was filled in for the deleted target word “to shatter,” for example). The texts were therefore considered to provide sufficient contextual support.

As can be seen, there is sufficient information provided for the reader to understand the nature of the task that these learners were being asked to do.

In addition to treatment materials, assessment materials may also appear in this section or, alternatively, this section may be divided into two sections, one dealing with treatment materials and another with testing/assessment

materials. An example of assessment materials from a study on think-alouds and reactivity is found below. The authors measured comprehension, intake, and written production following a think-aloud task. Only a portion of the description for each measure is provided (Leow & Morgan-Short, 2004, p. 45). In all three instances, the actual tool is provided in the appendix:

To measure participants' comprehension, an 11-item comprehension task was designed to elicit 17 pieces of information based exclusively on the advice, tips, or recommendations provided through the imperatives found in the text. The information was elicited predominantly via short and multiple-choice answers . . .

To measure participants' intake of the targeted forms, a multiple-choice recognition task was prepared. The 17 items on this task were also based exclusively on the advice, tips, or recommendations provided through the imperatives found in the text . . .

To measure participants' controlled written production of the targeted forms, a fill-in-the-blank task comprising 17 items that provided a list of advice for leading a healthy life was prepared . . .

The materials section presents a description of the actual materials used, but does not specify how they were used. The procedures section provides that information.

1.3.4.3 Procedures

The next questions that a reader can expect to be informed of include logistical issues related to what was actually done. How exactly was the task carried out? How was the treatment administered? How and when was testing done? The following is the procedures section from the study discussed above (Fukkink, Blok, & de Glopper, 2001, p. 482):

Participants were tested individually. Sessions started with a standardized explanation of directions to the students. It was decided that each text would first be read orally by the student, because reading aloud first appeared to encourage giving oral definitions in a pilot study and a previous study (Van Daalen-Kapteijns, Elshout-Mohr, & de Glopper, 2001). Students tried to decipher the meaning of the target

word thereafter in response to the question, “Which explanation does the dictionary give for this word?” Students were permitted to reread the text.

A warming-up task was introduced first, using materials that were similar to the experimental task. The experimental items were introduced only if students demonstrated adequate understanding of the procedure. The order of items was randomized for each participant. The sessions were tape recorded and transcribed for coding.

The Fukkink, Blok, and de Glopper (2001) study contained a separate section for scoring, in which detail was provided as to how responses were scored. A subsequent analysis section presented information about the statistical procedures used to analyze the data.

In some studies, authors present a visual image of the procedures followed, as in Figure 1.1.

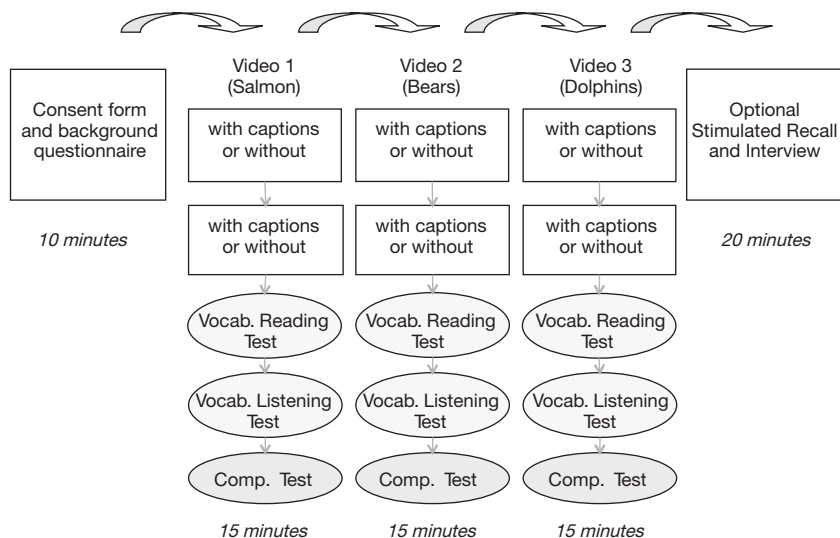


FIGURE 1.1 Example of graphic image of procedures

Source: Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology, 14*(1), 65–86, p. 71. Reprinted with permission of authors.

In other words, there are many ways to ensure that the reader understands what was done, how it was done, and the order in which various parts of an experiment were carried out.

1.3.4.4 Analysis

In some research reports, the mode of analysis may be a separate section or may be included in the results section. We present two examples of what might be included in a discussion of how one will analyze the results. The first, from Leow and Morgan-Short (2004, p. 46), provides information about the scoring procedure, and the second, a study on planning and narrative writing, from Ellis and Yuan (2004, p. 72), presents information about the statistical procedures to be used. In the first case, the section is called “Scoring Procedure” and in the second, the section is titled “Data Analysis.”

Scoring Procedure

For the recognition and controlled written production tasks, one point was awarded to each correct answer, and no points for incorrect answers, for a total of 17 points. The comprehension task was scored in the following manner: For all items except item 1, one point was awarded for each correct answer and zero for an incorrect one. For item 1, five out of seven correct responses were required before one point was awarded. For item 11, answers could have been provided in either English or Spanish.

Data Analysis

The normal distribution of the three groups' scores on all variables was tested in terms of skewness and kurtosis. A series of one-way ANOVAs were subsequently performed followed by post hoc Scheffé tests where appropriate (i.e., if the F score was statistically significant). In the one variable where normal distribution was not evident . . . , a Kruskal-Wallis Test was run, followed by independent *t*-tests to compare the pairs of groups. The alpha for achieving statistical significance was set at .05. Additionally, effect sizes were calculated . . . to examine the size of the effect of the different kinds of planning on performance of the task . . .

It is not always the case that all of these categories appear in every research report. Some may be combined, and others may not be relevant. The precise

organization of the report will depend on the design of the study and the authors' preference for presentation of the data.

1.3.5 Results

In this section of a research article, the results are presented with verbal descriptions of data, which are also often displayed in charts, figures, or tables. Results sections usually provide objective descriptions presented without interpretation. The excerpt below is a small part of a results section from Philp (2003, p. 110).

The provision of recasts depended entirely on the production of nontargetlike forms by each learner. Generally, as illustrated in Table 2, each learner received 44–55 recasts of question forms over five sessions with those in the Low group generally receiving higher numbers of recasts. Of these recasts, all groups received over 60 percent of recasts of stage 4 questions.

TABLE 2 Recasts provided to each group

Group	n	Recasts		Percentage of question forms in recasts		
		N	M	Stage 3	Stage 4	Stage 5
High	15	659	43.93	7 (44)	65 (415)	28 (179)
Intermediate	11	531	48.93	8 (42)	62 (316)	30 (155)
Low	7	379	54.14	6 (15)	63 (237)	33 (122)

As shown in Table 3, the High group was presented more frequently with long recasts (62 percent), whereas the Low group received more short recasts (67 percent). Similar numbers of recasts with one, two, or three or more changes to the learner's trigger utterance were received by all groups, although the Low group received slightly more of the latter. A comparison between groups is shown in Figure 1.

TABLE 3 Length of recasts and changes to learners' utterances in recasts: Percentages by group

Group	n	Length of recast		Number of changes		
		Short	Long	1 change	2 changes	≥ 3 changes
High	15	38	62	39	30	31
Intermediate	11	52	48	37	31	32
Low	7	67	33	30	30	40

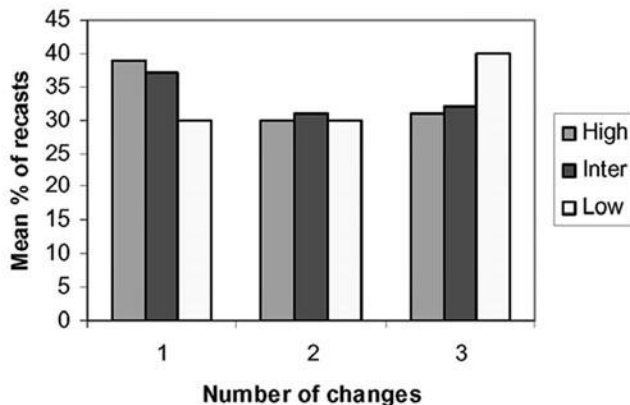


FIGURE 1 Comparison by group of proportion of number of changes in recasts

Philp, J. (2003). Constraints on "noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, 25, 110. Copyright © 2003 by Cambridge University Press. Reprinted with the permission of Cambridge University Press and with the permission of J. Philp.

Additional information about statistical results is also presented in the results section, as seen below (Philp, 2003, p. 111).

Results

To test hypothesis 1, which predicted that recall of recasts would be more accurate the higher the level of the learner, the High, Intermediate, and Low groups were compared. The results of a one-way ANOVA, provided in Table 4, show a significant effect for learner level on recall of recasts. With an alpha level of .05, the effect of learner level on recall of recasts was statistically significant, $F(2,30) = 4.1695$, $p < .05$. A priori contrasts, tested by the *statistic*, were computed to establish the source of difference between groups. A significant difference was found between the High and Intermediate groups on the one hand and the Low group on the other ($p < .05$). The High and Intermediate groups were not significantly different in performance on recall ($p = .814$) . . .

1.3.6 Discussion/Conclusion

The discussion and conclusion are often two separate sections and are primarily interpretive and explanatory in nature. The main idea of the study may be restated and the findings summarized. Then, the findings are interpreted in light of the research questions and an explanation is attempted (perhaps with regard to findings that were contrary to expectations). Below is an example from a discussion section on form-meaning mapping by native and nonnative speakers (Jiang, 2002, p. 624) where the author, in three separate paragraphs, provides a summary, an interpretation (along with problems), and a possible explanation. Below are the first sentences from each paragraph.

Summary statement

The results of experiment 1 show that whether an L2 word pair shares a single L1 translation does not affect native speakers' performance in the rating task . . .

Interpretation and problems

Although the findings of experiment 1 are consistent with the L1 lemma mediation hypothesis, there are two potential problems that have to be resolved before one can interpret the findings as evidence in support of it . . .

Explanation

One explanation for this discrepancy may lie in the possible involvement of conscious knowledge about L2 words in the rating task on the part of the nonnative speakers . . .

Finally, many studies include a section on the limitations of the study and suggest ways of remediating the limitations. Possible topics for future research may also be included. Typical contents of discussion, conclusion, and limitations sections are also discussed at length in Chapter 11, where we provide tips on writing and reporting research.

1.3.7 Notes

In some journals, any parenthetical material in an article is placed in footnotes at the bottom of the relevant page. In other journals, this material may appear as endnotes, where all the notes are collected together at the end of the article. In addition, there is generally an author's note, often including contact information,

information concerning prior presentations based on the research presented in the paper, and acknowledgments.

1.3.8 References

In most journals in the second language research field, everything cited in the paper appears in the reference list, and all sources listed in the reference list are cited in the paper. There is no single style used by all journals in the field; different journals have different styles for references. The use of style manuals is further discussed in Chapter 11.

1.3.9 Appendices

The appendices to a research article may include examples of the actual materials used in the study, along with any other information that, while necessary for the interpretation of the study, might interrupt the flow of the paper if included in the body of the article.

Now that we have provided a brief description of what can be expected in a typical quantitatively oriented article in the field of second language research, we move on to the main focus of this book, namely how to do second language research. We begin by considering the identification of research questions.

1.4 IDENTIFYING RESEARCH QUESTIONS

The first issue, and perhaps one of the most difficult aspects of any research undertaking, is the identification of appropriate research questions. Research questions are an integral part of quantitative research. The identification process for qualitative research, discussed in Chapter 7, is often quite different than it is for quantitative research. For example, in qualitative studies, in keeping with the goals of research, questions are often not as narrowly constrained.

Questions need to be interesting in the sense that they address current issues; at the same time, they need to be sufficiently narrow and constrained so that they can be answered. Broad questions can be difficult, if not impossible, to address without breaking them down into smaller answerable questions. For example, a general research question such as “What is the effect of the native language on the learning of a second or foreign language?” cannot be answered as formulated. This is because it represents a research area, but not a specific research question. To address the research area, a researcher might investigate the effect of a native language on specific aspects of a target language (e.g., phonology, syntax). One way to begin to reduce the general question would be

to consider the learning of a language that has a linguistic category not present in the native language. Again, this is somewhat broad, so the researcher might want to further reduce this to a specific question: “How do learners of a non-tonal language learn to make lexical distinctions using tone?” This is a reasonable starting point for the investigation of this question. The researcher could then examine the interlanguages of native speakers of English learning Chinese. Of course, the researcher would have to determine whether he or she wanted to examine production or comprehension in order to come up with specific hypotheses and a design. We return to the issue of hypotheses in section 1.4.2 of this chapter.

Where do research ideas come from? We mentioned earlier that research investigations need to be current, which of course entails that the questions have not already been answered in the literature, or have only partially been answered, and therefore require further or additional investigation. Research questions also need to be theoretically interesting; otherwise, we run into a “so what?” response to our research. Most reasoned research questions come from a reading of the literature and an understanding of the history of current issues. The conclusion sections of many articles suggest questions for future research. Some are quite specific, while others are merely suggestive.

TIME TO DO ...

Below are examples from three journals in which the authors refer to future research. Consider one or two of them: What research question could you develop from these suggestions?

1. A study of lexical repetition as a function of topic, cultural background, and development of writing ability by learners of English who are native speakers of Arabic, Japanese, Korean, and Spanish.
Future studies may wish to examine other possible topic-related variations, including distinctions between personal and nonpersonal writing and among different writing purposes. A second question is whether the time limitation imposed on these essays encouraged the use of repetition as a cohesion strategy (Reynolds, 2001, p. 472).
2. A study on the acquisition of English causatives by native speakers of Hindi-Urdu and Vietnamese.
There is, nonetheless, a need for further research in this area, involving a larger repertoire of verb classes, as well as a wider

range of proficiency levels . . . Similarly, further research could be undertaken on the influence of L1 verb serialization in languages like Vietnamese on the acquisition of the argument structure of verbs in nonserializing languages like English . . . Further research could also include studies on the acquisition of semantic classes relevant to various syntactic phenomena, involving a variety of languages (both as L1s and L2s), with different morphologies, classes of verbs and selectional restrictions on verbs (Helms-Park, 2001, p. 94).

3. A study of the relationship between speech and reading in a group of ESL learners who are native speakers of Japanese, Chinese, Korean, and Persian.

[F]urther exploration of the alphabet group differences on the figure pairs might prove productive. This should include a more detailed analysis of three groups (Roman alphabet, non-Roman alphabet, and Ideographic) instead of two groups as in the present study. In addition, the students might be presented with the decision tasks in their native language as a further control against the test effects . . . (Muchisky, 1983, pp. 94–95).

Another way to develop research questions is through the extensive reading and analysis of existing research. This can lead to the identification of gaps that may strike a reader as important. Often in reading an article, one might recognize that something has not been controlled for or that different languages might function differently on a certain important dimension. Alternatively, some controversy may have been left unresolved. This information may turn out to form the basis of a follow-up study, but a researcher must first make sure that others have not conducted such studies. A first step in this process is to consult a citation index (see your university librarian) to locate work that has cited the paper that you will be basing your study on. Another way of locating relevant information is through Web-based searches, which often yield studies published in a range of venues.

On other occasions, ideas for research might stem from observing learners either in or out of a classroom context or through some general feeling of curiosity having observed nonnative speaker linguistic behavior. These ideas may or may not develop into research studies, but in any case the first task is to do a literature search to see what has been done. There are many databases available for this purpose. Again, university librarians can assist with this process, and Web-based searches can often yield fruitful results.

TIME TO THINK ...

Return to the first “time to think” box in this chapter, in which you were asked to think about two to three questions about language learning and/or language teaching that you have been puzzled about.

Can you now take one or two of them and turn them into a research question?

1.4.1 Feasibility

The feasibility of a study may depend on a number of factors, some of which we have already mentioned (e.g., the breadth of the study in relation to its research questions’ scope and answerability). Another factor to take into account when considering feasibility is whether or not it will be possible to obtain the data necessary to answer the question. Consider a study in which one wants to conduct a survey of the attitudes of heritage learners (i.e., students who are learning the language of their parents, grandparents, etc.). In order to do this, the researcher first has to define exactly what constitutes a heritage learner. One question might be whether someone can be considered a heritage learner if he or she has distant relatives in Uzbekistan, for example, but has only very rarely heard the language spoken. Or, is a heritage learner limited to a family where the language was regularly spoken in the home while the language of the environment was the language of schooling and all outside of the home transactions? Following this step, the researcher needs to go about identifying individuals who would qualify under the definition chosen. In many settings, it would be difficult to find a reasonable number of participants to make the study interesting. Thus, pertinent data sources need to be identified as a part of determining the feasibility of the study.

Another study might seek to compare performance on different communication task types. As we discuss in Chapter 3, there are many important dimensions on which communicative tasks can differ. However, it might not be feasible to require participants to do 15 different tasks. Exhaustion and boredom might set in, and the researcher would not know how to interpret the results. This is not to say that such a study could not be conducted; it is just that the design of the study might entail large numbers of participants who may or may not be available for the many rounds of data collection that such a study would necessitate.

Thus, any study should be designed with a full understanding of the fact that the limitations of the setting and the population might constrain the research.

1.4.2 Research Questions and Hypotheses

Research problems are generally expressed in terms of research questions and/or hypotheses. Research questions are the questions for which answers are being sought, while research hypotheses can be used to express what the researcher expects the results of an investigation to be. The hypotheses are based on observations or on what the literature suggests the answers might be. There may be times when, because of a lack of relevant literature, hypotheses cannot be generated because the researcher is dealing with new and/or unexplored areas.

The literature review that leads up to the hypotheses should report all sides of an issue. In other words, fair and complete reporting is essential in any research study. We return to the issue of hypotheses in Chapter 5.

To see examples of both research questions and hypotheses, consider the following from Lakshmanan (1989, pp. 84–86). This was a study that investigated the acquisition of verb inflection and the use of pronouns by children learning English (native speakers of Spanish, French, and Japanese). The data, collected by other researchers, are from longitudinal studies of these three children. Below are five research questions from this study.

- RQ#1. Do null subjects in the interlanguage (IL) of these child L2 learners decrease with time?²²
- RQ#2. Is there a developmental relation between null subjects and verb inflections in the IL of these child L2 learners? In other words, is increase in verb inflections accompanied by a corresponding decrease in the use of null subjects?
- RQ#3. Are obligatory verb inflections acquired at the same time for all the categories of verb morphology or does the acquisition of verb inflections depend on the specific category of verb morphology (e.g., be copula, auxiliaries be, do, have, present 3rd singular regular, past regular etc.)?
- RQ#4. Is there a developmental relation between null subjects in is constructions (is copula and auxiliary utterances) and is constructions? In other words, does the proportion of null subjects present in is contexts increase with the increase in the proportion of is constructions?
- RQ#5. Are there any differences between the distribution of null subjects and subjects in is constructions and non-is constructions in these child L2 learners' IL?

As can be seen, these research questions are expressed as explorations of relationships. Lakshmanan also formulated them as hypotheses. Examples of hypotheses stemming from the above research questions are given below.

- Hypothesis 1. Null subjects in the four subjects' IL will decrease with time.
- Hypothesis 2. There is a negative relation between the development of verb inflections and the use of null subjects; in other words, null subjects will decrease with the increase in verb inflections.
- Hypothesis 3. The acquisition of obligatory verb inflections depends on the specific category of verb morphology.
- Hypothesis 4. There is a positive relationship between the use of null subjects in is constructions and the development of is constructions.
- Hypothesis 5. There are significant differences between the distribution of null subjects and lexically realized subjects in is constructions and non-is constructions. The frequency of occurrence of null subjects will be greater than the frequency of occurrence of lexically realized subjects in is contexts; the frequency of occurrence of null subjects in non-is contexts will be less than the frequency of occurrence of lexically realized subjects in non-is contexts.

In other instances, the research questions are expressed as one item, as in Révész, Sachs, and Hama (2014, p. 623), and with different degrees of elaboration. In the third of their questions, the results of previous research were mixed and, in the last of their research questions, there was no previous research to base any prediction on.

- 1) Do different intended levels of task complexity result in different levels of cognitive load during task performance, as measured by expert judgments, performance on a secondary task, and eye tracking?

It was hypothesized that the task version designed to be more complex would impose greater cognitive load.

- 2) What is the effect of task complexity on the acquisition of the English past counterfactual construction under task conditions where recasts are provided?

Based on the Cognition Hypothesis, which drove our operationalization of task complexity along resource-directing dimensions, it was hypothesized that learners who received recasts during tasks with greater reasoning demands would show more development.

3) What is the effect of manipulating the input frequencies of verbs (skewed vs. balanced) on the acquisition of the English past counterfactual construction under task conditions where recasts are provided?

To date empirical results have been mixed regarding the effects of input frequency manipulations; thus, the null hypothesis was assumed.

4) Is there an interaction between the effects of task complexity and input frequency in the acquisition of the past counterfactual construction under task conditions where recasts are provided?

To our knowledge, this is the first empirical study to address this question. Thus, the null hypothesis was assumed.

The above examples have been expressed in terms of yes/no questions, or as questions that have a dichotomous possibility as an outcome, but as research areas become more nuanced, it is often the case that we need to dig deeper. In other words, we know whether something has an effect on something else; what we need to explore is how great the effect is. For example, Dale, Harlaar, and Plomin (2012), in a study of early development of twins, point out that because of previous research, they are able to address a wider range of questions. These are posed as “to what extent” questions rather than “yes/no” questions, as is seen below (p. 32), particularly in questions 1, 4, and 5:

- (1) How well do measures of early and adolescent L1 learning, and adolescent reading, predict SLA, both individually and collectively?
- (2) How does the etiology of L2 achievement compare with that of early and adolescent L1 and adolescent reading?
- (3) Is there evidence for differential etiology of poor or exceptionally capable L2 achievement relative to the rest of the normal distribution of L2 achievement?
- (4) To what extent do the same genes account for variability in L1 and L2 achievement?
- (5) To what extent do the same environmental influences account for variability in L1 and L2 achievement?

Nassaji (2012), in his study investigating teachers’ perceptions of the relationship between SLA and language teaching, asked the following questions (p. 344):

- (1) To what extent are teachers familiar with SLA research?
- (2) How easily can they access SLA research, and what sources do they consult?

- (3) To what extent do they read research articles and, if they do not read them, what are their reasons?
- (4) How do they judge the relevance and usefulness of SLA research for classroom teaching?
- (5) How do they perceive the relationship between researchers and teachers?
- (6) What are their expectations of SLA research?

The important point to remember is that defining research questions requires an understanding of previous research. This is so because we need to know what an interesting question is in relation to what we already know and we need to further previous knowledge by extending that knowledge with questions that suggest a more continuous answer (to what extent) rather than just a *yes* or a *no* answer.

TIME TO THINK ...

Given the questions in Nassaji's study, what type of data do you think would be appropriate to answer his questions (e.g., questionnaires, observations, experimental)?

1.4.3 Replication

Replication is a central part of the development of any field of inquiry. If one cannot repeat the results of a particular study, the validity of the results of the original study might be called into question.³ In fact, the *Publication Manual of the American Psychological Association, Sixth Edition* states, "The essence of the scientific method involves observations that can be repeated and verified by others" (American Psychological Association, 2010, p. 12). Similarly, Albert Valdman, the then editor of the journal *Studies in Second Language Acquisition*, has asserted that "the way to more valid and reliable SLA research is through replication" (1993, p. 505). As Porte (2002) further notes, without critical replication studies, "it will be extremely difficult ever to discover the definitive response to a research question or hypothesis found in one particular study . . . which then permits us to generalize those findings to fit exactly another context of language learning" (p. 35). Porte (2012) defines a replication study (p. 3) as one that "attempts to discover whether the same findings are obtained by another research in another context, and whether the outcome appears to reflect knowledge which can therefore be separated from the context in which it was originally found." It is thus crucial that researchers report in enough detail to allow others to determine with precision what has been done. Unfortunately,

since much research in the field of second language learning is published in journals, space constraints often preclude full and complete reporting. To this end, Polio and Gass (1997) recommend that researchers submit detailed appendices for publishers to keep either online or as hard copies if journal space is limited, although publishers have not yet fully embraced this idea. More specifically, Polio and Gass suggest that these appendices include information about any guidelines used for coding the data, measures of proficiency or development, instruments for data elicitation (including pre- and post-tests), experimental protocols, and biodata on the participants. Generally speaking, there are two primary reasons for replication: verification and generalizability. We return to the issue of replication in Chapter 11.

1.5 CONCLUSION

In this chapter, we have dealt with some of the basics of L2 research, including the range of different types of research that exist, what to expect from a typical research report, and how to identify research questions, generate hypotheses, and consider issues such as feasibility and the role of replication in second language research. In Chapter 2, we deal with the question of research ethics, focusing on the important issue of informed consent.

POINTS TO REMEMBER

- Research is a systematic approach to finding answers to a question.
- There are many approaches to conducting research.
- Two main orientations are quantitative and qualitative.
- There are standard components to a quantitatively oriented research report.
- A first step in conducting research is to identify a research question.
- Some considerations in developing a research question include ensuring that it addresses current research and that it is feasible to conduct research to answer the question.
- Replication and full reporting are important parts of the research process.

MORE TO DO AND MORE TO THINK ABOUT ...

1. Consider the journals you found from your search in section 1.1. Can you determine the scope of each journal? What kinds of topics do they deal with? Some journals are quite explicit; others might require a look through the tables of contents and abstracts.
2. Select three of these journals and consider the extent to which the articles follow the framework set up in this chapter. If they do not, in what way(s) do they deviate?
3. Consider these same journals. Do the journals give guidelines for submission (e.g., length, style guidelines, number of copies to submit, mode of submission)? List the guidelines you have found.
4. Find three articles and consider the end of the discussion section or perhaps the beginning of the conclusion section to determine if the authors acknowledge limitations of the study. What do they say?
5. Consider the following research questions that are expressed as yes/no questions:
 - a. Is motivation important for second language learning?
 - b. Is it important to like your teacher to be successful at learning a second language?
 - c. Is learning a language by living in the second language environment as effective as learning a language in the classroom?

Assume that previous research has already answered them in the affirmative. Change them into questions that provide more detailed information by asking them not as dichotomous yes/no questions, but ones that suggest a more continuous nature of the response.

6. Read the conclusion sections from three different articles in three different journals. Do the authors point to future research possibilities? If not, do they do this elsewhere (perhaps shortly before the conclusion)? What do they say, and are there any issues that are of interest to you?
7. How can the following research topics be turned into researchable questions?

Example:

Gender differences in language classes

Do males perform differently than females on a grammar test following treatment in which negative feedback is given?

- a. Motivation
- b. Task effectiveness
- c. Novice teacher performance
- d. Attention
- e. Final grades

NOTES

1. According to the *Publication Manual of the American Psychological Association, Sixth Edition*, the word *participant* is more appropriate than words such as *subject*. In a section titled “Guidelines for reducing bias,” they state, “although descript terms such as *college students, children, or respondents* provide precise information about the individuals taking part in a research project, the more general terms *participants* and *subjects* are also in common usage” (American Psychological Association, 2010, p. 73). They go on to say that when discussing statistical results, the word *subjects* is appropriate (e.g., *between subjects design*).
2. Null subjects refer to expressions in languages such as Italian or Spanish that have verbs with no overt subjects. In Italian, for example, to say *I speak Italian*, one can say *Parlo italiano*, where the first word means *I speak*. The overt word for *I (io)* is not used.
3. Along with the issue of replication comes the important issue of data reporting. How much should be reported? How much detail? The simple answer is: enough so that someone can replicate the study.