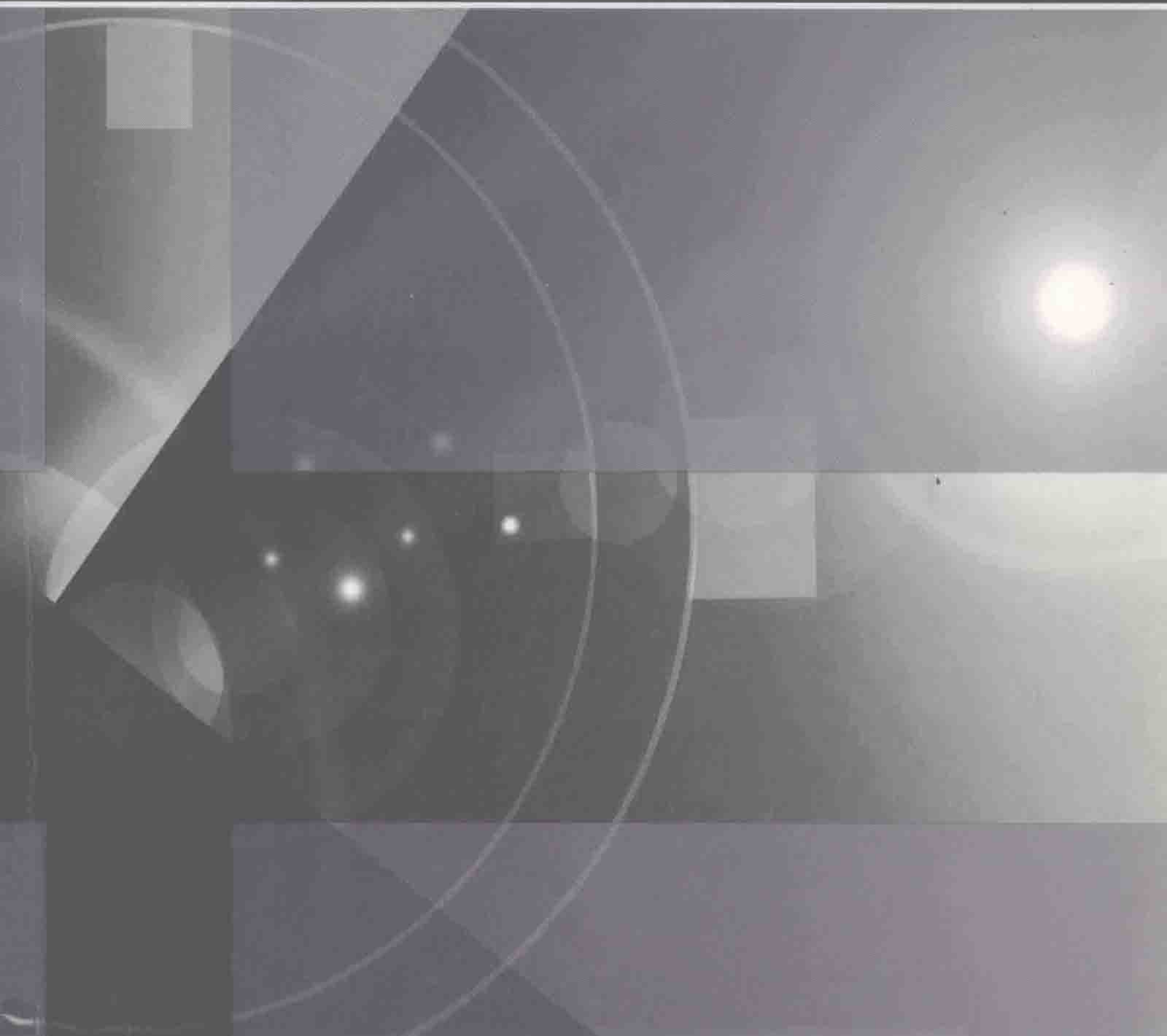
Data Elicitation for Second and Foreign Language Research



Susan M. Gass • Alison Mackey

DATA ELICITATION FOR SECOND AND FOREIGN LANGUAGE RESEARCH

Susan Gass
Michigan State University

Alison Mackey
Georgetown University

LEA
2007 LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey London

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430
www.erlbaum.com

Cover design by Tomai Maridou

Library of Congress Cataloging-in-Publication Data

Gass, Susan M.
Data elicitation for second and foreign language research / Susan M. Gass, Alison Mackey.
p. cm.

Includes bibliographical references and index.
ISBN 978-0-8058-6034-4 — 0-8058-6034-7 (cloth)
ISBN 978-0-8058-6035-1 — 0-8058-6035-5 (pbk.)
ISBN 978-1-4106-1628-9 — 1-4106-1628-2 (e book)
1. Language and languages—Study and teaching—Research—Methodology. 2. Second language acquisition—Research—Methodology. I. Mackey, Alison. II. Title
P53.755.G37 2007
418.007'2—dc22 2006024276
CIP

Books published by Lawrence Erlbaum Associates are printed on acid-free paper, and their bindings are chosen for strength and durability.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

Second Language Acquisition Research
Theoretical and Methodological Issues
Susan M. Gass, Jacquelyn Schachter, and Alison Mackey, Editors

---

Schachter/Gass • *Second Language Classroom Research: Issues and Opportunities* (1996)

Birdsong • *Second Language Acquisition and the Critical Period Hypothesis* (1999)

Ohta • *Second Language Acquisition Processes in the Classroom: Learning Japanese* (2001)

Major • *Foreign Accent: Ontogeny and Phylogeny of Second Language Phonology* (2001)

VanPatten • *Processing Instruction: Theory, Research, and Commentary* (2003)

VanPatten/Williams/Rott/Overstreet • *Form-Meaning Connections in Second Language Acquisition* (2004)

Bardovi-Harlig/Hartford • *Interlanguage Pragmatics: Exploring Institutional Talk* (2005)

Dörnyei • *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition* (2005)

Long • *Problems in SLA* (2007)

VanPatten/Williams • *Theories in Second Language Acquisition* (2007)

Monographs on Research Methodology

Tarone/Gass/Cohen • *Research Methodology in Second Language Acquisition* (1994)

Yule • *Referential Communication Tasks* (1997)

Gass/Mackey • *Data Elicitation for Second and Foreign Language Research* (2007)

Gass/Mackey • *Stimulated Recall Methodology in Second Language Research* (2000)

Markee • *Conversation Analysis* (2000)

Dörnyei • *Questionnaires in Second Language Research: Construction, Administration, and Processing* (2002)

Of Related Interest

Gass • *Input, Interaction, and the Second Language Learner* (1997)

Gass/Sorace/Selinker • *Second Language Learning Data Analysis, Second Edition* (1998)

Gass/Selinker • *Second Language Acquisition: An Introductory Course, Second Edition* (2001)

Mackey/Gass • *Second Language Research: Methodology and Design* (2005)

For more information on LEA's Second Language Acquisition Research Series please contact us at www.erlbaum.com

Preface

In the process of writing our 2005 volume, *Second Language Research: Methodology and Design*, one chapter stood out: Chapter Three, entitled "Common Data Collection Measures." Although we included a great deal in that chapter, we also felt that there was much more that could be said. Sure enough, even though that chapter ended up being the longest in the book, upon reading it, many of our colleagues, co-authors, students, and editors expressed a desire to know more, asking us questions like "what about discussing the problems with this" and "I wish you'd included an example of that." Not only for them, but also for ourselves, we took up the challenge of demystifying the choice and creation of data elicitation techniques. The current book, devoted to data collection, is therefore an extension and expansion of this earlier chapter and, for those who want to delve further into the actual collection of data, can be used as a companion to our 2005 text.

One of the most difficult tasks when conducting second language research is identifying and deciding on ways to collect data. There are intricate relationships between research questions and data elicitation techniques, and with so many possibilities, it can be hard to keep all of the necessary considerations straight. Fortunately, a body of practical knowledge has been built up in this area, and we hope that this book takes a step toward synthesizing and systematizing it, thereby helping L2 researchers to be inventive without having to reinvent the wheel. In many graduate programs, students in disciplines such as applied linguistics, second language studies, TESOL, psycholinguistics, cognitive science, and educational linguistics are expected to collect data for projects, graduate seminar work, theses, and dissertations. Although many programs offer classes in research methods and design, there is clearly a need for a textbook that provides step-by-step examples and specific guidelines on how to collect second language data.

In this book we begin each chapter by discussing common research questions underlying data elicitation measures in that area, we then go on to identify relevant types of data according to a continuum that includes naturalistic language, prompted linguistic production, and non-linguistic experimental responses. Whenever possible, we provide the history and origins for some of the techniques, and we also provide detailed descriptions in terms of "how to's," summaries of published studies that have made use of the techniques described, and discussions not only of advantages, but also of common misconceptions, caveats, and/or limitations related to various techniques. Interested readers can then seek out the studies we have summarized, both for further details and to better understand the usefulness and relevance of the technique to their own research. We hope that our book will stimulate researchers to use data elicitation measures with an understanding of their strengths and limitations and, when appropriate, to modify and combine techniques to match their own research needs.

There are many individuals to whom thanks are due. First, we are grateful to our editor, Cathleen Petree of Lawrence Erlbaum Associates, for her (as usual) unwavering support for this project and for obtaining extremely useful external reviews of both the proposal and the final product. Both Susan Gass and Alison Mackey had research assistants, and we are indebted to them for their help in many ways. At Michigan State University, Robin Roots showed her usual keen ability to pinpoint problems and suggest alternative wording. Both Robin and Junkyu Lee found and summarized many relevant articles. At Georgetown University, Rebecca Sachs, in particular, contributed enormously to this effort, and served as an extremely talented editor. Jaemyung Goo, Mika Hama, Rebekha Abbuhl, and Bo-Ram Suh also assisted in important ways. All of our assistants provided library work and advice, criticisms, and occasionally even praise. Their careful work and contributions greatly improved the final version of this text. We are also grateful to Jane Ozanich and Jenefer Philp, who graciously contributed some of the drawings (those on p. 114 are the work of Jane Ozanich, those on p. 117 are by Jenefer Philp, and those on pp. 113 and 125 are by Jamie Lepore Wright). We appreciate their willingness to share their artistic talents with us. Finally, we sincerely thank the aforementioned colleagues and students who encouraged us to write this book. We have learned a great deal ourselves in the process and hope that it helps to answer some of the questions they had.

Contents

CHAPTER 1

Introduction

How Do Second Language Researchers Decide on Data Elicitation Measures?

When we read research reports in journals or books, we can easily be lulled into a false sense of security. The straightforwardness and simplicity of the reporting, while often necessary for publication, can obscure the trials, tribulations, difficulties, and, in some cases, significant problems that were part of the research process. The passage from generating an idea for a research question to publishing a report is rarely tidy or obvious; it is a long and arduous undertaking. Furthermore, since we cannot control every eventuality, compromises in data collection and analysis are often necessary. These sorts of details usually are not disclosed in the reports that we read. To give just one example, because of the attrition of participants, it is quite common for researchers to have to eliminate variables or factors that are potentially of interest. A lack of participants or data in a study can force researchers to narrow their focus or to separate a large study into smaller studies, dealing with separate variables each time. For clarity and parsimony in reporting, however, these developments may be glossed over in the final report. Although there are many potential problems along the road for researchers to be aware of when conducting studies, this book deals in detail with perhaps the most significant one of all: the collection of the data.

Second language research relies heavily on empirical and often experimental studies that involve the interpretation of data. The term *second language* is used in this book as a cover term to include, *inter alia*: (1) learning in both second and foreign language environments, (2) learning in both classroom and naturalistic settings, and (3) learning a second language as well as learning a third, fourth, language, and so on. The terms *acquisition* and *learning* are used interchangeably, meaning only that non-primary language acquisition is the focus.

One of the most difficult tasks when conducting research in such a broad field is identifying and deciding on methods of data collection. In recent years, to answer an increasing variety of research questions, the field has witnessed a burgeoning of techniques to elicit data. At times, however, the relationship between the techniques of elicitation and the questions being asked has been obscured. In this book, our goal is to demystify the process of data collection. We present suggestions for collecting data in relation to a wide range of research areas and, in so doing, provide a detailed picture of the many different techniques available for eliciting data on the knowledge and use of language, intuitions about language, and/or attitudes toward language. Data collection in second language research is clearly a multi-dimensional topic. This book reflects that fact by exploring some of the common measures that have appeared in the literature over the years and across a range of domains of SLA research.

In Mackey and Gass (2005), we pointed out that data elicitation techniques are limited only by one's imagination. For expository purposes in the current text, we have divided the book into chapters according to research area, contextualizing elicitation techniques by linking each to a specific area of second language research. We opted to begin with the brain in areas of cognition, move to memory, mind and language and then move to interaction and sociolinguistic context and end with the classroom. Each chapter deals separately with data collection for studies based on psycholinguistics (processing), cognitive processes, capacities and strategies, linguistics, interactions, sociolinguistics and pragmatics, surveys, and classrooms. Within each chapter we have organized the data collection methods according to a continuum ranging from naturalistic data, prompted production data to prompted response data—a classification scheme based on Chaudron's (2003) exposition of data collection methods. Whereas some research questions appear to call for eliciting and measuring language learners' spontaneous speech (naturalistic data), for others it may seem more appropriate to use formal tools to collect written or spoken language production (prompted production data) or to elicit specific responses to prompts (prompted response data).

At the same time, an important consideration when using any elicitation instrument is to understand its advantages and limitations, as well as to realize that many techniques are versatile enough to shift easily between research areas and elicit data to address a wide range of questions. Thus, in contextualizing each technique within a specific area of research, it is not our intent to suggest that these techniques are limited to those research areas (and vice versa), but simply to illustrate common practice. Indeed, creative uses and combinations of techniques from a range of areas are often appropriate and necessary both to triangulate data and to move research forward. As will be discussed in more detail below, no measure is foolproof, and each must be tried before actually using it for collecting data. In this book, where appropriate, we also provide some historical context, discussing the origins of the data collection techniques presented. Finally, for illustrative purposes, we present descriptions of published second language studies in boxed insets, as well as samples of data collected using the different techniques.

1.1. PILOT TESTING

Before delving further into the details of data elicitation techniques, a caveat: Regardless of how carefully researchers have designed their studies, they should never simply assume that their data collection methods will work. A crucial issue that must be kept in mind in the data-gathering process is the importance of pilot testing. A pilot study is generally considered to be a small-scale trial of the proposed procedures, materials, methods, and (sometimes) coding sheets and analytic choices of a research project. The point of carrying out a pilot study is to finalize these essential components, to uncover any problems, and to address them before the main study is carried out. A pilot study is therefore an important means of assessing the feasibility and usefulness of the data sampling and collection methods and revising them before they are used with the research participants. While it might seem that careful prior planning would allow researchers to skip this step, it is in fact critical to allocate additional time to conduct pilot tests, as they can reveal subtle flaws in the design or implementation of the study—flaws that may not be readily apparent in the plan itself, but that could otherwise prove costly and time-consuming, perhaps even leading to the loss of valuable and irreplaceable data.

Since pilot studies sometimes result in data that might be useable for the main study, it is important to note the constraints imposed by human subjects committees or institutional review boards (also referred to as ethics review boards, discussed extensively in Mackey & Gass, 2005), which must approve research conducted in most university or institutional settings. Some researchers choose to seek permission from their institutional review boards to carry out their experiments in such a way that if they do not encounter problems with their pilot-testing, they can use those data for their main study as long as exactly the same procedures are used. However, not all institutions will give blanket permission for this, and many will not consider requests for the retroactive use of data. It is therefore worthwhile to investigate these issues while also keeping in mind that it is a rare pilot study that does not result in some sort of revision of materials or methods.

1.2. THE SIGNIFICANCE OF DATA COLLECTION MEASURES

One goal of second language research is to uncover information about learner behavior or learner knowledge independent of the context of data collection. However, research findings are highly dependent on the data collection measures used. As such, although many research domains have common measures associated with them, and while the choice of a measure may be related to the theoretical framework within which the research is conducted, there should be no single prescribed elicitation method for a given domain, nor is there necessarily a "right" or "wrong" elicitation method for a given context. At the same time, saying that numerous elicitation measures can be used in the various areas of second language research does not imply that one method is as good as another; rather, the choice of one method over another is highly dependent on the research question being asked.

Also important to consider in data collection is the concept of construct validity, which refers to the degree to which the phenomenon of interest is adequately captured. It is often difficult to capture a particular construct, such as proficiency, aptitude, or motivation, with one measure. In cases such as these, construct validity can be enhanced when multiple estimates of a construct are used. For example, if we consider the construct "amount of input" and its relationship to some part of acquisition, construct validity might be enhanced if a variety of factors, such as length of residence, amount of language instruction, and the language used in the participants' formal education, are considered.

We mentioned above that research questions can dictate, to a certain extent, a researcher's choice of data collection method. Let us consider what this means by looking at some hypothetical examples. In each of these examples, we refer to specific elicitation techniques, each of which is dealt with in this book and appears in the subject index.

Researching a Grammatical Structure

After years of teaching Japanese, you recognize that English-speaking learners of Japanese have great difficulty with passives. You perform an extensive literature review and find out that there are theoretical reasons for this, so you decide to investigate the issue further. The task in front of you is to gather data from learners of Japanese to determine exactly which faulty generalizations they may be making. In other words, which language forms are used by learners at various stages of Japanese proficiency? Once you have determined the type(s) of data you need to elicit (i.e., samples of Japanese passives), your next task is to determine how to elicit the appropriate data.

Your first thought, before pilot testing, is to have learners describe pictures that depict various actions (e.g., a man being hit by a ball, a dog being kissed by a boy). Unfortunately, the learners, who are experts at avoidance, produce very few examples of the structure in question. You then modify the task and tell the learners to start with the object of the action. You even point to the man and the dog in the pictures. This does not work very well, either, and the learners do not do what you have asked. You are thus left with the question: Did they not produce the requisite passive because (1) they do not have the linguistic knowledge to do so, (2) the active sentence is easier to formulate, or (3) they did not understand how to carry out the task? Only the first interpretation will help you in dealing with your research questions but there are too many possibilities for you to be able to interpret the data legitimately. It is therefore necessary to question the value of this elicitation method.

You then realize that you have to "force" the issue and make the learners behave in a way that allows you to be confident that you are obtaining information that reflects their actual knowledge about passives. There are a few ways that this is commonly done in second language research. One way is to use what are known as acceptability / grammaticality judgments, in which a list of grammatical and ungrammatical sentences is presented to learners who are then asked to indicate whether they consider them acceptable Japanese sentences or not. This is followed by a request to the learners to correct those sentences they have judged to be incorrect. Forcing correction allows researchers to ensure not only that their target of investigation (e.g., passives) is included in the sample of sentences, but also that learners are indeed focusing on passives. Another way to gather information about passives is through "elicited imitation." In this method, sentences are read to learners (usually in the form of audio recordings to ensure that everyone hears the same sentences at an identical rate and with identical intonation), and the learners are asked to repeat them. As with acceptability judgments, researchers can control all aspects of the sample sentences. A third possibility for eliciting information about passives is known as "truth-value judgments," which might be particularly useful in the case of Japanese passives because some of the differences in language forms involve subtle meaning differences. With truth-value judgments, learners are given short contextualized passages with relevant sentences embedded in them. Following each passage is a question designed to ascertain whether or not the learners can correctly interpret the meaning of the passage.

Thus, the investigation of a particular grammatical structure offers a number of possibilities for data elicitation measures, the choice of which will depend on the questions being asked (e.g., acceptability judgments or elicited imitations if researchers wish to gather information about grammatical knowledge, or truth-value judgments if the focus is on subtle meaning differences). In any event, specific research questions can be used to narrow the choice of data collection measures.

Interaction Research

> Suppose your research interest is to find out whether recasts or negotiation will lead to faster development of relative clauses in a second language.[1] You find groups of English learners of Italian who are at four different stages of development in their knowledge of relative clauses. For each group, half of them will serve in your "recast" group and the other half in your "negotiation" group. You first give a pre-test to ensure comparability of groups. Everyone then completes a picture-description task in which feedback is provided by a native speaker of Italian, either in the form of recasts or negotiation. At the end of the session, you give the learners a post-test on relative clauses and then a delayed post-test three weeks later. You find that there are no differences between the groups. When you go back to analyze the actual transcripts of the sessions, however, you realize that your findings are probably due to a lack of examples of relative clauses in the data.

This example illustrates how important it is in task-based research first to ascertain whether or not tasks will, in fact, elicit the targeted grammatical structures and provide opportunities for interactional feedback. Since relative clauses are frequently used to differentiate one object or person from others (e.g., *the boy who is wearing a red hat is my brother*—in other words, not the boy who is wearing the green hat), researchers targeting relative clauses must make sure that their tasks are consistent with the normal function of relative

---

[1]Recasts occur when an incorrect non-native speaker (NNS) utterance is repeated in a more target-like way while maintaining the original meaning.

Example:	NNS:	I have three bird my picture.
	NS:	You have three birds in your picture?

Negotiation occurs when learners and their interlocutors try to understand the meaning or form of an utterance.

Example:	NNS:	I see bud my picture.
	NS:	You see what in your picture?
	NNS:	I see bud my picture.
	NS:	Bud?
	NNS:	Yes
	NS:	Do you mean bird?
	NNS:	Ah, bird, yes.

clauses and involve learners identifying an object or person from among others. In sum, researchers need to ensure (through piloting) that each elicitation measure yields the kind of data that will be useful in addressing their research questions.

Researching Pragmatics

Assume that you want to conduct research on the pragmatic problems that a particular group of students might have (e.g., English speakers learning Chinese). You have further limited your area of research to interactions between English learners of Mandarin and their Mandarin-speaking professors. You obtain permission to observe interactions between these two groups of people in order to determine what sorts of pragmatic errors may occur, but after five days of observations, you have little in the way of consistent results. Why might that be the case? One reason might be that you have not sufficiently narrowed down your research question. For example, you might be considering too many pragmatic functions (e.g., complaining, apologizing, requesting, inviting) rather than constraining the data. A second reason is that waiting for language events such as these to occur often depends on luck. You might end up with some interesting examples that could be fodder for insightful qualitative analyses, but if you are looking for sufficient examples to be able to make quantitative generalizations, you may need to force the issue through written discourse completion tests or well-designed role plays. In their simplest form, discourse completion tests present learners with specific contexts in which responses are required. Role plays involve acting out situations and are also useful for establishing particular contexts. An elicitation method could involve a learner and a researcher sitting at a table when the researcher pretends to accidentally knock over a glass of water. What the learner actually says to accept the apology could then be recorded.

As can be seen through the examples above, research questions can help to guide researchers in selecting appropriate elicitation measures. In conducting research, researchers must understand how data elicitation methods relate not only to their general areas of research, but also to their more specific research questions.

1.3. ELICITATION PITFALLS

As we begin a research project, the first question that confronts us is the identification of a research area; the second task is to narrow the area down into