

METHODS FOR STUDYING LANGUAGE PRODUCTION



EDITED BY

LISE MENN

NAN BERNSTEIN RATNER



METHODS FOR STUDYING LANGUAGE PRODUCTION

Edited by

Lise Menn

University of Colorado

Nan Bernstein Ratner

University of Maryland



2000

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS

Mahwah, New Jersey

London

Copyright ©2000 by Lawrence Erlbaum Associates, Inc.

All rights reserved. No part of this book may be reproduced in any form, by photostat, microfilm, retrieval system, or any other means, without the prior written permission of the publisher.

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430-2262

Library of Congress Cataloging-in-Publication Data

Methods for studying language production / edited by Lise Menn, Nan Bernstein Ratner

p. cm.

Includes bibliographical references and index.

ISBN 0-8058-3033-2 (c : alk. paper). — ISBN 0-8058-3034-0

(p : alk. paper)

1. Language acquisition—Research—Methodology. 2. Language disorders—Research—Methodology. I. Menn, Lise. II. Ratner, Nan Bernstein.

P118.M47 1999

401'.43'072—dc21

99-12768

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

Contents

1	In the Beginning Was the Wug: Forty Years of Language-Elicitation Studies <i>Nan Bernstein Ratner and Lise Menn</i>	1
I.	ELICITING KNOWLEDGE OF LANGUAGE	
2	Infants' Verbal Imitation and Their Language Development: Controversies, Techniques, and Consequences <i>Elise Frank Masur</i>	27
3	Examining Young Children's Morphosyntactic Development Through Elicited Production <i>LouAnn Gerken</i>	45
4	Coining New Words: Old and New Word Forms for New Meanings <i>Eve V. Clark</i>	53
5	Children's Innovative Verbs Versus Nouns: Structured Elicitations and Spontaneous Coinages <i>Ruth A. Berman</i>	69
6	Methods for Studying the Production of Argument Structure in Children and Adults <i>Jess Gropen</i>	95
7	Methods for Stimulating and Measuring Lexical and Syntactic Advances: Why Fiffins and Lobsters Can Tag Along With Other Recast Friends <i>Keith E. Nelson</i>	115
8	Eliciting Second Language Speech Data <i>Catherine Doughty and Michael H. Long</i>	149
		V

II. GATHERING PRODUCTION DATA IN NATURALISTIC SETTINGS

- 9 What You See is What You Get: The Importance of Transcription
for Interpreting Children's Morphosyntactic Development
Carolyn E. Johnson 181
- 10 Food for Thought: Dinner Table as a Context for Observing
Parent-Child Discourse
*Barbara Alexander Pan, Rivka Y. Perlmann,
and Catherine E. Snow* 205
- 11 Exploring Register Knowledge: The Value
of "Controlled Improvisation"
Elaine Andersen 225
- 12 The Story Behind the Story: Gathering Narrative Data From Children
Richard Ely, Anne Wolf, Allyssa McCabe, and Gigliana Melzi 249
- 13 Studying Conversation: How to Get Natural Peer Interaction
Susan M. Ervin-Tripp 271

III. DEVELOPMENTAL DISORDERS

- 14 Elicited Imitation and Other Methods for the Analysis
of Trade-Offs Between Speech and Language Skills in Children
Nan Bernstein Ratner 291
- 15 The Challenge of Studying Language Development
in Children With Autism
Helen Tager-Flusberg 313
- 16 Understanding Grammatical Deficits in Children With Specific
Language Impairment: The Evaluation of Productivity
Laurence B. Leonard 333
- 17 Influences of School-Age Children's Beliefs and Goals on Their
Elicited Pragmatic Performance: Lessons Learned From Kissing the
Blarney Stone
Mavis L. Donahue 353

IV. ADULT DISORDERS

- 18 Jean Berko Gleason's Contributions to Aphasia Research:
Pioneering Elicitation Techniques
Harold Goodglass 371
- 19 Studying the Pragmatic Microstructure of Aphasic
and Normal Speech: An Experimental Approach
Lise Menn 377
- 20 Eliciting Language From Patients With Alzheimer's Disease
Loraine K. Obler and Susan De Santi 403
- Index 417

1

In the Beginning Was the Wug: Forty Years of Language-Elicitation Studies

Nan Bernstein Ratner
University of Maryland

Lise Menn
University of Colorado

All fundamental scientific innovation must marry new ways of thinking with better styles of seeing. Neither abstract theorizing nor meticulous observation can provoke a change of such magnitude all by itself.

—Gould (1998, p. 18)

This book represents some major approaches to collecting language production data from children and young adults, but it is not intended as a complete handbook, which would require a much larger volume. The use of elicited narrative in cross-linguistic comparative developmental study, for example, has been richly elaborated elsewhere, especially by Berman and Slobin (1994), and is not covered here; for narrative in cross-linguistic aphasiology, Menn and Obler (1990) may be useful. Other aspects of production and methods of eliciting specific syntactic forms have been covered in the first several chapters of McDaniel, McKee, and Cairns (1998). For priming studies, see Bock, Lobell, & Morey (1992) and other

papers cited in Bock and Levelt (1994). The CHILDES data analysis (see also Sokolov & Snow, 1994) and the CHILDES Bib database can be searched for more information on how database corpora have been used (www/childes.psy.cmu.edu).

This book also is limited in its treatment of cross-cultural issues in data collection: The way people talk to children in front of strangers or while they are being recorded may differ a great deal across cultures, and even more when the stranger is a foreigner. This said, we are delighted to present a wider range of topics in language elicitation than has ever been brought together. All the contributors speak from experience, and we have all learned the art of elicitation by apprenticeship and by trial and error. We have created this book in the hopes of reducing the number of trials that you, the reader, will have to endure.

LOOKING BACKWARD: IN THE BEGINNING WAS THE WUG

Before the late 1950s language production was studied naturalistically, with rare efforts to elicit language knowledge through experimental paradigms. This was particularly true for the study of child language (Bar-Adon & Leopold, 1971). We begin our book on methods of studying language production, begins with a review of the paradigm that changed the nature of developmental psycholinguistic research: Berko Gleason's (1958)¹ wug test. Through the use of hand-drawn figures of whimsical creatures performing odd acts, Berko elicited plurals, past tenses, and diminutives, among other morphological affixes, by presenting the now classic task: "This is a picture of a wug. Now I have two of them. I have two ____."

The responses of Berko Gleason's subjects' provided the first experimental evidence that young children had productive knowledge of the morphological patterns of their language, generalizable to novel word forms. Prideaux (1985) noted:

¹ For comprehension of spoken language, Brown's (1957) article on children's comprehension of the morphological markers of syntactic category ("some niss," "a niss," "nissing") seems to be the first experimental study.

Bogoyavlenskiy (1957/1973), translated and reprinted in Ferguson & Slobin (1973), independently developed a production paradigm with three stimuli that he used to test eight children. However, this research was not available to most students of child language acquisition until 15 years after Berko's study made its mark on the field.

Jean Berko, in her pioneering work in language acquisition, set the stage for a great deal of experimental work by providing both conceptual and methodological orientations which have proven very valuable and productive . . . [this] work constitute[d] a landmark in language acquisition research. It set the stage for a series of studies employing her basic methodology, or variations of it, which have reaped rich rewards over the . . . years. (pp. 8–9)

Other researchers called the paradigm “ingenious” (Clark & Clark, 1977; Maratsos, 1979) and “a great methodological contribution” (Slobin, 1973) to child language acquisition research. Uniquely, the wug was both a controlled and naturalistic method for eliciting language behavior, at least for children who are used to playing utterance-completion games with their caregivers.

The wug has been used in many ways over the years. A citation search of published research that has utilized “wug like tasks,” as they are sometimes called, reveals extensions of the technique almost immediately following its original publication. Many researchers sought to replicate Berko’s findings of morphological knowledge by young children, extending the original scope of inquiry to speakers of other languages and bilinguals,² children with varying diagnoses of specific language impairment or delay,³ children and adults with mental retardation,⁴ reading-impaired children,⁵ and children from differing socioeconomic and dialect communities.⁶ Even a cursory listing of such efforts reveals their scope and remarkable frequency.

Subsequent work with English-speaking children was able to detail more finely the progress that children make toward mastery of the individual morphemes targeted by the wug task (Anisfeld & Tucker, 1967; Derwing & Baker, 1979; Ervin, 1964; Menn & MacWhinney, 1984; see also Prideaux, 1985, for discussion of unpublished research using the wug technique). Refinement of the design also allowed researchers to evaluate the adequacy of competing accounts of how regular affixation is mastered (see Derwing & Baker, 1986).

² For example, a representative listing would include the following: French: Hiriartborde & de la Bouillerie (1973); Spanish: Perez Pereira (1989); Russian: Zakharova (1973); Hungarian and German: MacWhinney (1978); Finnish: Lyytinen (1987); Quechua: Isbell (1972); and Bilinguals: Gray and Cameron (1980).

³ For example, Wiig, Semel, & Crouse (1967), and Bellaire, Plante, & Swisher (1994).

⁴ For example, Dever (1972) and Dever & Gardner (1970).

⁵ For example, Wood (1985) and Mahoney (1994).

⁶ For example, Ramer & Rees (1973).

Almost immediately following publication of Berko Gleason's research, clinicians noted the possible extension of the wug paradigm to formal assessment of language performance in children with suspected disorder. The Berry-Talbott Test was but one of many attempts to adapt the wug task as a tool for assessment (see Berry, 1969; more recently, Channell & Ford, 1991; Rubin, Patterson, & Kantor, 1991). However, clinical use of the wug task has been sporadic because of concerns about its relation to spontaneous language performance and the appropriateness of its normative expectations. More recently, as researchers focus attention on familial forms of specific language impairment (SLI; e.g., Gopnik & Crago, 1991), the "wug test" has been used to assess the nature of the underlying deficit in this population (Goad & Rebellati, 1994). Not only has nonsense-word affixation been tested, but the very ability of SLI children to associate the novel word *wug* with a variety of concepts has been appraised under varying levels of support (see Ellis Weismer & Hesketh, 1993). Similarly, the wug test has been used to assess conceptual generalization in Down syndrome children (Duffy & Wishart, 1994).

A review of citation lists of Berko Gleason's (1958) *Word* article over the years gives an interesting mini-view of the evolution of developmental psycholinguistics. References to the wug task appear among the first issues of many seminal journals of our time, such as *Applied Psycholinguistics* and the *Journal of Child Language*. In the first 15 years following publication, the article was extensively cited by researchers attempting to validate its utility and extend its findings to nontypical populations. Over time, however, it became apparent that children from infinitely varied backgrounds can in fact generate appropriate morphological inflections for words they have never encountered before; thus, the fact that almost any human being can do the task (albeit some more successfully than others) became much less interesting than the question of how it is accomplished.

A number of alternative hypotheses can explain how children perform on wuglike tasks (see Prideaux, 1985, for a particularly thorough discussion). The nonsense-word paradigm continues to feed active debate on the adequacy of competing models in explaining morphological acquisition and the potential role of modularity in language learning and use (cf. Prasada & Pinker, 1993); thus, after a period of relative quiet, citations of the article burgeoned in the 1990s. The wug has become a sort of rallying point for theories and simulations of how networks learn to generalize to novel events (e.g., Rumelhart & McClelland, 1986) and whether, for instance, the ability of children to inflect is best interpreted by connectionist models or by symbolic, rule-based architectures. The original inter-

pretation of the wug studies, that children do not learn to inflect by rote or mimicry, has been tempered over the years by proposals that some language learning may indeed occur by rote, as well as by analogy and rule (MacWhinney, 1978). As theory now grapples with the results of functional brain imaging tasks, the wug task surfaces in discussions of how best to interpret neuro-imaging data on past tense processing (e.g., the debate between Jaeger et al., 1996, and Seidenberg & Hoeffner, 1998).

Because this is a book about language-elicitation techniques, some analysis of the strengths and weaknesses of the historic technique is appropriate. Extended application of the paradigm brought with it meta-analysis of its utility, including thoughtful discussion of the role that Berko Gleason's clever original pictured stimuli (or any pictured prompts) played in supporting children's performance on morphological tasks, as well as the inevitable discussion of whether nonsense-word affixation tasks were legitimate windows into the process by which children learn morphology or are able to demonstrate their competence (cf. Levy, 1987; Lewis & Windsor, 1996). It has become increasingly clear that subjects of all ages and degrees of linguistic sophistication tend, for the most part, to do more poorly on nonsense-word affixation tasks than on those that utilize real words of the language (e.g., Miller & Ervin, 1964; Perez Pereira, 1989); in fact, Berko herself noted this discrepancy in comparing the responses of her young subjects to the nonsense words and selected real-word stimuli. As thesis advisor to the original study, Brown (1973) provided a careful analysis of how to relate Berko Gleason's data to patterns of affixation in spontaneous language. He suggested important differences between controlled elicitation of grammatical affixes and spontaneous language use, including the child's need to understand cue words in the experimental carrier phrases (e.g., "yesterday" and "whose") and ability to focus on relevant aspects of the supporting pictorial stimuli.

Brown noted that Adam, Eve and Sarah's spontaneous language performance did not pattern exactly like performance by older children in the original wug study. His 1973 text spends more than 10 pages considering whether these differences can be accounted for by differing criteria for judging morpheme productivity. An immediate observation was that Brown's famous "14 morphemes" list collapsed all allomorphic variations into single morpheme classes (such as plural or past tense), whereas Berko Gleason's careful sampling of the varying morphophonemic contexts elicited a wide range of performance from children that seemed to depress accurate production of a supposedly early acquired morpheme seen in spontaneous language. In an analysis similar to one that motivated Bryant

and Anisfeld's (1969) experimental inquiry into morphological performance, Brown noted that there are probably at least three levels of difficulty in the original wug-set stimuli: those stem + affix pairs that only require the child to know a basic form of affix and to automatically apply voicing assimilation (dogs, cats; tagged, walked); those that pose an additional burden of specifying the voicing (days; played); and those that are phonologically context and morpheme specific (houses; patted). Although obligatory contexts for the more difficult forms are less frequent in children's spontaneous speech samples than are those for the first group of affixes, Berko Gleason sampled broadly and somewhat evenly across the allomorphs, thereby depressing the children's apparent level of competence—as well as producing differences in patterns of acquisition between her data and Brown's subsequent data. In a discussion that foreshadowed later concern over the roles that imitation and rule abstraction might play in language development (cf. MacWhinney, 1978; Pinker, 1991), Brown also speculated that some real-word plurals might, after all, be stored as wholes, whereas others might not.

Brown also noted another important methodological difference between spontaneous language sampling and elicitation tasks that is relevant to many concepts discussed later in this volume. Language sampling tends to isolate patterns in language from large amounts of data produced by small numbers of subjects, whereas Berko Gleason's study paved the way for eliciting data from large numbers of subjects. Cross-sectional patterns of failure or success of individual children may thus fail to duplicate patterns of mastery that emerge during the development of a single child. Finally, the novelty of the task itself may induce changes in children's tactics for language use that do not mirror their tendencies in natural conversation. This has also been a chronic concern of researchers over the years in comparing structured vs. naturalistic means of data collection, and numerous contributors to this volume address the concept. After analyzing the data in a number of ways, Brown (1973) was confident that Berko Gleason's data could be reconciled with the order of acquisition data that flowed from analysis of his own Harvard children. He closed his discussion with the observation that naturalistic and elicited methods "together give us the best chance of discovering the truth" (p. 293). Many of the chapters in this volume agree with his assessment, and they show how researchers have tried to marry structured questions about language with more naturalistic contexts for their exploration.

Other researchers noted limitations in extrapolating children's productive knowledge of inflectional morphology from wuglike tasks. Levy (1987) suggested that older children may handle the task differently from

younger children, who are still acquiring the basics of their language system; her youngest subjects showed closest parallels between the nonsense-word task and affixation of real nouns. Dever (1973) noted a poor correlation between the characteristics of spontaneous speech and performance on wuglike tasks by children with mental retardation, suggesting that the technique has less utility as a predictor of spontaneous language performance in this population. Indeed, the fact that many children with cognitive impairment find the task immensely difficult sheds interesting light on the nature of their generalization difficulties (Ratner, 1998).

The wug task is most often referenced in discussions of the acquisition of inflectional morphology. However, the 1958 nonsense-word task was also used to study the abilities of children to use derivational morphemes. Children in the original study found these morphemes surprisingly difficult to use, and this observation paved the way for numerous other studies in the same area (cf. Derwing & Baker, 1986). It is also sometimes easy to forget that the wug paradigm was only part of the original set of tasks that Berko used to explore children's morphological development. In the 1958 study, children were also asked to speculate on the origins of common English compound words; interest in children's capacity to coin and de-compose such terms continues to be of interest to researchers today (see Clark, chap. 4, this volume).

The passage of time has left its mark on the wug. During the 1970s, the genderless wug found itself included in a critique that found too few female characters and role models in speech and language tests (Rabe & Matlin, 1978). Like many visual icons that have endured (e.g., cartoon characters or corporate logos), the wug has undergone some minor facelifts over time (see Fig. 1.1). Although (like many of us) it seems to have put on a little weight over the years, it seems to be busily teaching future generations of students of psycholinguistics. Perusal of textbooks and a web search of psycholinguistics and language acquisition syllabi show an almost universal use of assignments designed to encourage (or require) that students test the paradigm for themselves. (In one interesting



Fig. 1.1. The wug parade. From left to right; Berko Gleason (1958); reprint of the article in Saporta (1961); Berko Gleason & Ratner (1998).

version, the instructor has “morphed” the wug into Henry Moore–like statuary, a tribute to the artistry of the original.) As an enduring concept in psycholinguistic research, the wug has become a generic, like Kleenex or Xerox, a concept so basic to what we know and do that increasingly it appears in the popular literature without attribution to its origins (cf. Pinker, 1994; and press discussion by Allman, 1991, and Safire, 1994).

In any field, the discovery of facts is of paramount importance in moving knowledge forward. However, the discovery of techniques is no less important: We can only find keys where there is light to search by. Perhaps no innovation other than the invention of the tape recorder has had such an indelible effect on the field of child language research.

Theories come and go, including those that account for how a child knows that more than one *wug* is *wugs*. Although, we are still learning from Berko’s insight that children can be asked to count cats or dogs, we can learn even more, and be challenged much more, if we also ask them to count wugs—these oddly charming little birds and their whimsical friends.

METHOD, DATA, AND THEORY: FORTY YEARS ON THE ROAD

What method should one use to investigate a particular problem? There is, of course, the issue of whether to study production at all; many psycholinguists privately admit that they prefer to study comprehension, because production data are “too messy.” Well, yes, they are messy. But messes can yield surprises, which pose challenges, which can lead to reformulations of issues and whole new ways of looking at things. That’s the kind of excitement some of us live for.

How do methods of data collection shape researchers’ view of language knowledge and behavior? Of many possible examples, the following are two recent findings for consideration.

1. Verb-argument structures: Roland and Jurafsky (1998) showed that the frequency of occurrence of particular verb-argument structures is systematically different depending on whether one analyzes a corpus of connected text or requests subjects to make individual sentences using specified verbs.
2. Lexicon, syntax, pragmatics: As people who study child language corpora have known for years, after children’s mean length of utterance (MLU) reaches about 3.5, it no longer is much of an

index of syntactic development; instead, it is highly dependent on what the children are talking about and what role the children are playing in the conversation. An extreme example is provided by research at the University of Colorado (Biasca, 1999): Narratives elicited when a language-delayed pre-reader pretended to read a story book yielded low-frequency vocabulary ("queen," "witch") and longer formulaic elements ("they have to save Snow White because you have to go on your hat top of your head") appropriate to the genre, linked together syntactically and paratactically but with virtually no semantic or pragmatic coherence.

The list of such interactions between task and performance is endless, because the experimenter's (or observer's) method itself inevitably generates unintended effects, a type of noise that is called *experimental artifact*. Such noise can be so overwhelming as to invalidate the argument; the entire data set may then turn out to be as deeply flawed as the idea that the earth is flat.

As researchers, we all (at least secretly) hope that our data—if not our conclusions—will be able to stand throughout our own lifetimes, if not longer. If we have done a good enough job of controlling for interfering factors, this may be the case; later researchers may be able to use what we have done, and the field as a whole may show an orderly progression of successive approximations to understanding—that is, an early researcher's conclusions will still stand as roughly correct, and the "noise" in their data will become the data for later, more refined work. In that case, we have the delicious paradox of scientific research: The wilderness apparently resolves into the promised land, which, on closer examination, turns out to be full of new doubts and questions, and so on, world without end.

What can one do to make this happy spiral toward truth more likely to occur in one's own work? Take method seriously. Consider how theory shapes method. It is theory, explicit or implicit, that determines what observed events are to be regarded as data (the "good stuff" from which legitimate conclusions can be drawn) and what observed events are to be discarded as noise (random, irrelevant, and interfering with the good stuff). Method is a third and equal partner with theory and data; they are mutually dependent and deserve equal scrutiny. Theory is implicit in method, like it or not; and whether your data are really capable of supporting your theory depends on the validity of the procedure you used to obtain them. This does not make theory a villain: On the contrary, it is not

really possible to recognize experimental (or observational) artifacts for what they are until a theory that will make them visible has started to emerge!

In this volume, chapters by Johnson (chap. 9) and Donahue (chap. 17) raise the issue of how method affects the kind of data that are available for analysis (see the following section of this chapter). Johnson's chapter deals with the ways that orthographic transcription can distort early morphological data; Donahue's chapter shows that children's longer-term social agendas affect their pragmatic behavior in an experimental interactive task.

Pragmatic theory has made it possible to see the artifacts of some classic methods. For example, Tanz (1980, 1983) documented a variety of pragmatic artifacts in her own and other studies of semantic development. Her principal technique was to ask a child to ask a third person a question; she then examined what the child said, looking for immature language errors and also for what appeared to be mistakes in the child's interpretation of her instructions. One persistent kind of interpretation error that her subjects made caused Tanz to reinterpret some of the results of one of the most famous early works on language development, Chomsky's (1969) *The Acquisition of Syntax in Children from 5 to 10*. In the paperback edition of that work, the cover design shows two comic strip-style speaker "balloons." One balloon says, "Ask him who this is." The other balloon says, "Mickey Mouse." This epitomizes one of Chomsky's generally accepted conclusions: children, even after age 5, may interpret both "ask" and "tell" as meaning "tell."

Tanz replicated this result, first accidentally (Tanz, 1980) and then by deliberate experimental manipulation (Tanz, 1983) of pragmatic conditions. She found that a number of children interpreted "ask" as either "ask" or "tell" according to pragmatic considerations. If they didn't know the answer, they asked the question. If they did know or if they had a stake in a particular answer, they showed a tendency to provide it. This suggests that children who tell when requested to ask may not have a problem with the semantics of "ask," after all. Rather, instead of interpreting instructions to ask someone something as a request to be carried out literally, they may interpret them as a sincere request for information; being cooperative, they simply go ahead and supply the information if they know it.

Other widely accepted conclusions about syntactic and semantic development have been reevaluated in recent years and have been shown to be based on experimental designs that induced pragmatic artifacts (see Foster-Cohen's, 1994, study of binding theory, which concludes that

“pragmatic factors must be systematically controlled in any evaluation of syntactic knowledge” [p. 237], and Clark & Grossman’s 1998, reexamination of the role of “mutual exclusivity” in early word learning, which is discussed in chap. 17, this volume). Other famous developmental conclusions may also be ripe for pragmatically based reconsideration. For example, Donahue discusses Siegal’s (1991) demonstration that the classical administration of the Piagetian number conservation task (“Are there more in this [closely spaced] row or this [widely spaced] row?”) is contaminated by pragmatic artifacts that systematically underestimate the age at which number conservation is achieved. One is reminded of Clark’s (1992) citation of a sign that he said was posted for years in a London hospital: “No head injury is too trivial to ignore.” (Read it again!)⁷

Moral: When it is (or seems to be) obvious what something ought to mean, bottom-up processing of the syntax and semantics is likely to be short-circuited or overruled by pragmatics, in production as well as in comprehension. After all, someone *wrote* that sign!

A LOOK AHEAD

The chapters that follow include paradigms old and new, approaches reconsidered, re-configured, and refined to gain increasing insight into the complexities of language acquisition and use. They range across the age span, across levels of human ability, and across contexts in which language is used and learned. Some of the paradigms described here involve one experimenter and simple stimuli; others involve complex situational prompts; still others cannot be contained in the laboratory or clinic but rely on observation of language in its natural context of use.

Some contributors to this volume follow Berko Gleason in using novel or invented linguistic forms, or they prompt children to create them (Berman, chap. 5; Clark, chap. 4; Gropen, chap. 6; Nelson, chap. 7). Others eschew the question of what speakers might do to investigate what speakers actually do, in the real and messy laboratory called life (Ely, Wolf, McCabe, & Melzi, chap. 12; Ervin-Tripp, chap. 13). From the phoneme to the conversation, these authors explore the most fruitful methods for obtaining data and analyzing language use. They range from

⁷ In case that sentence still seems to make sense to you, suppose it said “No head injury is too trivial to treat.” That means “Every head injury is important and should be treated.” So “No head injury is too trivial to ignore” should mean “Every head injury is important and should be ignored.”

a focus on the implications of narrow transcription (Johnson, chap. 9) to the broadest contexts of conversational interaction (Andersen, chap. 11; Ervin-Tripp, chap. 13). Johnson points out that (just as in a pointillist painting) the whole is dependent on the smallest of parts: Our very concepts of children's grammatical development are inextricably linked to the adequacy of our representation of their speech attempts.

Cross-linguistic testing of research hypotheses is also necessary to the construction of adequate theories of language. From discussion of acquisition patterns in languages other than English, to disordered language use, to second language acquisition, the contributors tackle the sometimes difficult considerations of how to translate (a) questions originally asked about English development and use into (b) equivalent questions about languages of considerably different structure (Andersen, chap. 11; Berman, chap. 5; Clark, chap. 4; Leonard, chap. 16; Menn, chap. 19).

The main axis used to organize the chapters of this book is the degree of control that the researcher attempts to exert over what the person being studied will say. This could be considered, conversely, as the degree to which the language produced is intended to be representative of natural speech; for as control loosens, naturalness increases. The tightest control and the most contrived language is in elicited imitation; the loosest control and the most natural language occur in "fly-on-the-wall" studies of people deeply involved in conversation with one another. The more natural a study is, the more it tells you about what people do say (under the circumstances of the particular interaction); but the less it may tell you about what people can say if they need to. A natural study in July will find no evidence for the knowledge of Christmas carols; a July study that elicits Christmas carols will find that people can indeed sing them, but it will completely lack information as to how, where, and why they are sung.

The naturalness-control continuum has been understood and exploited reasonably systematically since Labov's sociolinguistic work of the late 1960s (Labov, 1972), and presumably all researchers understand that there is no single right method for studying language production. The only issue is whether the results of the method a researcher plans to use give results that are a valid basis for the type of argument he or she wants to make. As noted earlier, even the wug task is not immune to pragmatic bias and the kind of light it sheds on the child's linguistic knowledge. But as researchers we can only wish that despite our own methodological shortcomings, our work will survive the test of time as well as the wug has.