1. Introduction

Language is intangible: it results as an interaction of high-level abstract cognitive abilities within human beings. What aspects of human cognition are we seeing through the narrow window of linguistic data? In any research endeavor targeting an aspect of language, researchers cannot avoid the issue of what exactly is manifested in the data they collect and analyze. Consequently, methodology is not an issue of concern only for psycholinguists.

Probably the most uncontroversial method in linguistic research would be collecting and analyzing of naturalistic data (e.g. a “natural” conversation) with minimal control or intervention by the researcher. As discussed by Sonja Eisenbeiss in Chapter 1 of the current volume, naturalistic data provides material with “high ecological validity” (p. 12) (i.e. uncontrolled and hence natural) to analyze. However, it is very difficult for linguists to compare the patterns observed between different participants in such studies; observations made with naturalistic data are muddled by the influence of contextual and communicative factors, among other cognitive and physical conditions of the participants. Furthermore, expressions missing from a corpus do not necessarily indicate that they are not a part of the speakers’ linguistic ability: those expressions might be missing simply because there was no communicative reason to use them. For these reasons, naïve analyses made using naturalistic data are not extensive enough for one to consider what the knowledge of language consists of.

Experimental methods with adequate control of the subjects, tasks, and
interpretation of the data are necessary to address the issues raised in theoretical linguistics. Chomsky (1957: 16), for example, argued for the use of context-free grammatical judgment by native speakers of the language as a legitimate source of data which a linguist can consider to test his/her hypothesis, using the well-known pair of \([\text{Colorless green ideas sleep furiously}]\) and \([\text{Furiously sleep ideas green colorless}]\). In reality, however, as Schütze (1996) points out, “Unlike natural scientists, linguists are not trained in methods for getting reliable data and determining which of two conflicting data reports is more reliable.” (p. 4) The insufficient training in research methods has received more direct attention in the field of psycholinguistics, where participants often are young children or “laypeople” that do not have the expertise or training that professional linguists are expected to bring to the data.1

Methodological considerations are particularly important in the tradition of children’s language acquisition studies, as young children are sensitive to pragmatic and meta-linguistic influences and hence are easily biased by carelessly designed experiments, as shown in the criticisms presented by Hamburger and Crain (1982). As more psycholinguistic experimental studies contributed to the research of linguistic theory (e.g. L1 acquisition studies inspired by the Principles and Parameter Approach in Chomsky (1986)), psycholinguists realized the need to provide reference material for researchers who are interested in the acquisition of syntax but have not received systematic training in methodology. McDaniel et al. (1996) successfully introduce readers to a wide range of experimental methods used with extensive references to actual studies. Crain and Thornton (1998) focus on the methods to study young children’s syntactic knowledge in the context of advancing the understanding of Universal Grammar (UG). Since UG is considered to be innately available according to generative grammar, studies of children’s L1 acquisition maintained a close connection to these theoretical investigations. Those books mostly contain “off-line” tasks (e.g. asking participants to answer questions, make judgments, and produce sentences, all after being presented stories or contextual settings), which do not require any specialized (and often pricy) equipment or training for the researchers.

However, during the following two decades, “on-line” task techniques (e.g. recording participants’ responses mechanically as they process linguis-

1 See Schütze’s (1996) discussion in chapter 4, where he carefully reviews the research concerning linguists’ and non-linguists’ grammatical judgments.
tic stimuli presented) with child participants have developed rapidly and become more prevalent, as the equipment has become more affordable and accessible. At the same time, advancements in psychological and linguistic theories of the knowledge of language have come to require more sophisticated empirical data, which directly provides insights into the issues of how researchers interpret observable linguistic behaviors obtained in experimental studies.2

The current volume can be taken as an updated successor of McDaniel et al. (1996). As the editors acknowledge (p. 3), they have reduced the discussion of off-line tasks significantly and included more on-line tasks and computational modeling, as well as testing methods used in second language acquisition. In the following section, an overview of the volume is presented, followed by a review of its strengths and limitations. Finally, there is a brief discussion of those who would benefit from this book.

2. Overview of the Volume

One characteristic of the current volume is the breadth of the tasks and research topics covered. The content is based on lectures given for a workshop held since 2003 at Utrecht University and the editors’ intention is to make the references useful to researchers of language acquisition who work in different fields with different frameworks. The chapters are divided into two parts (Part I and II).

In Part I, the first three chapters deal with off-line tasks, in which the participants produce linguistic data after the context is established or necessary information is presented to them. In Chapter 1, Sonja Eisenbeiss describes three types of production methods in language acquisition research studies: naturalistic production, semi-structured elicitation, and production experiments. Although Eisenbeiss acknowledges the versatility and “high ecological validity” (i.e. minimized influence by the research process or

2 In fact, there have been recent works which cast doubt on the language acquisition research which generative linguists often refer to as the empirical evidence for the innateness of UG. For example, Zukowski et al. (2011) argue for the need to reconsider the empirical data in Thornton (1990), which claimed that children are adult-like in their comprehension of the wanna-construction. In addition, Elbourne (2005) and Conroy et al. (2009) pointed out possible experimental bias in results of previous works which were argued to demonstrate young English-speaking children’s knowledge of Condition B (including Chien and Wexler’s seminal work in 1990).
tasks) of naturalistic studies, she also points out the problems caused by the lack of researcher control.

Discussion by Cristina Schmitt and Karen Miller follows in Chapter 2, which includes three comprehension methods: TVJT (Truth-Value Judgment Task), the Act-Out task, and the picture matching task (including the Magic Box task, developed to circumvent the problems pointed out in previous studies). They emphasized the importance of “triangulation,” or using all three methods in the same study to reduce any task-related bias. They warn the readers about excessively relying on the TVJT: “Although, it is assumed in the literature (Crain & Thornton (1998)) that the TVJT always taps into more than a preference, this is not always the case. When there are two possible interpretations, depending on details of the set-up, adults may or may not be biased towards one of the interpretations, which of course does not mean that they have an impaired grammar.” (p. 37)

Reflecting the recent development of methodology using comparative judgment (Schütze (1996)), in Chapter 3, Antonella Sorace provides a description of Magnitude Estimation in developmental linguistic research and tips for using it. Sorace emphasizes that using the n-point grading scale helps to capture empirical data, particularly when they are subject to gradience and optionality. This method can easily be adopted to collect monolingual native speakers’ grammatical judgments and hence should deserve the serious attention of researchers in theoretical linguistics.

In Chapter 4, Elizabeth Johnston and Tania Zamuner focus on tasks developed for infant and toddler testing. Tasks covered are the visual fixation procedure, the headturn preference procedure, and the preferential looking procedure. Even though the chapter could have been enhanced with the inclusion of photos or diagrams, readers will be able to gain a reasonable understanding of the factors that make it challenging for researchers to obtain empirical data from very young participants.

Chapters 5–7 describe various on-line processing methods. In Chapter 5, Judith Rispens and Evelien Krikhaar provide a well-organized introduction to using ERP (Event-Related Potentials) with examples of auditory, semantic and syntactic processing studies. It is emphasized that ERP is effective in investigating the time course of language processing in the brain. Chapter 6 by Julie C. Sedivy is an introduction to the use of eyetracking (eye movements). Although the same task is discussed in more detail in two chapters in Sekerina et al. (2008), Sedivy provides a useful comparison of the advantages and disadvantages of using eyetracking compared to off-line tasks. Theodoros Marinis introduces various other on-line processing meth-
ods in Chapter 7: word-monitoring, self-paced reading/listening, cross-modal priming, as well the self-paced listening with picture verification, which has recently been developed by Marinis. Marinis argues that on-line tasks are unconscious and automatic, plus they create fewer demands on working memory and are immune to problems caused by interference from metalinguistic factors.

Chapter 8 deals with a research method which has not been included in previous publications, and hence can be considered as a unique feature of the current volume. Lisa Pearl gives a compact and accessible introduction to using computational modeling in language acquisition studies. The chapter contains the rationale for using a probabilistic learning program, using algorithms, such as linear reward-penalty, neural networks, and Bayesian updating. At the end of the chapter, Pearl states: “... we must consider if the details of the model are psychologically plausible by looking at the data available on human language acquisition from theoretical and experimental research. We should remember that modeling is an additional tool we use to understand language acquisition, not a replacement for others we already have.” (p. 181)

Part I concludes with another unique chapter by Jan H. Hulstijn (Chapter 9) covering the testing of second language proficiency. After reviewing the issue of defining language proficiency, Hulstijn considers characteristics of different test constructions including vocabulary tests, sentence imitation, the cloze test, and elicited production (which can supplement the discussion of semi-structured elicitation in section 3.1 in Chapter 1). This chapter could be taken as a quick reference for testing in L2 research.

Part II consists of three chapters about comparing across groups and focuses on the factors that need to be considered as a researcher recruits and compares various subject groups for a study. Chapter 10 by Sharon Unsworth and Elma Blom compares three populations: L1 children, L2 children and L2 adults. With ample references, they discuss several issues: the age effect (the critical period), what to compare and at what ages to compare language proficiency, and how to determine the age of first exposure. They also include sections on controlling for cognitive development, proficiency levels (by using standardized tests, which relates to the content of Ch. 9), and input. Johanne Paradis discusses which issues need to be considered for a comparison of TD (typically-developing) children and children with SLI (Specific Language Impairment) in Chapter 11. Though this chapter could be taken as an updated version of a chapter by Laurence Leonard in McDaniel et al. (1996), Paradis includes a more detailed discus-
sion of the use of MLU (Mean Length of Utterance) to match subjects, as well as a discussion of the three-group design (TD sequential bilinguals to TD monolinguals and to monolinguals with SLI). Chapter 12 by Anne Baker and Beppie van den Bogaerde covers the issue of measuring the linguistic development in deaf learners. Throughout their review of theoretical and methodological issues, they emphasize, “... it is very difficult to find a homogeneous population of deaf children.” (p. 253) The chapter touches on the importance of the experimental setting’s formality, as well as the issue of transcribing sign language data.

Chapter 13 by Hugo Quené presents basic information about formulating and testing a hypothesis for an experimental study, which makes this chapter highly useful to advanced undergraduate or graduate students, or anyone who might be unfamiliar with experimental studies. Quené makes this chapter more accessible to researchers who are concerned with the statistical aspect of data collection by including FAQs about statistics, such as: “How many participants and items are required?” “How can I prove that there is no difference?” “What to do about missing data?”

In the following sections, an evaluation of the current volume will be given, followed by a discussion of how it could benefit readers from different backgrounds.

3. Evaluation

The most notable characteristic of the current volume is the consistency of the chapter formatting. Virtually all chapters are structured in the same format (Rationale, Linguistic Variables, Subjects, Description of Procedure, Analysis and Outcomes, Advantages and Disadvantages, Do’s and Don’ts). This format consistency should make it easier for readers to compare and understand the differences among methods. As the editors admit, the breadth of topics received priority over topic depth. Since each chapter is relatively short, readers who are not necessarily familiar with experimental methods will find this volume more accessible as well as being useful as a handy reference. Readers can make a direct comparison of off-line and on-line tasks without referring to several different publications. For example, Sedivy (Ch. 6) notes, “… [in off-line tasks] only the final response made by the child is considered,” (p. 133) as well as, “The [eyetracking] method is potentially costly in terms of human labor and equipment, and some considerable time is required for researchers to become acquainted with procedures for programming experiments for automated eyetrackers, or learning to code
and analyze the data.” (p. 133) Though this volume is not sufficient to make one prepared to conduct any of the tasks discussed, it provides an efficient overview of what experimental tasks are available for psycholinguistic research, with discussion of their strengths and limitations. The current volume also helps the reader realize that L1 or L2 research topics could be investigated using similar experimental methods.

Considering the broad coverage of experimental tasks, it is highly understandable that many authors of the chapters encourage including more than one method in a research project. This issue has been raised in “comprehension-production discrepancy” (Snyder (2007)), which suggests the possibility that naturalistic and experimental data are actually capturing different aspects of the linguistic abilities of children. In typical language acquisition studies, one experimental method is often employed; however, the use of multiple data-collection methods would minimize the risk of bias caused by any specific task and enable the researcher to get a more complete view of the nature of the participants’ linguistic knowledge.

It should also be noted that the editors provide a general discussion in the introduction about ethics, planning, recruitment, materials and equipment, and test sessions. As experimental studies are more widely used in the field of linguistics, the need for ethical education for researchers and students is also increasing. At the end of her detailed introduction to conducting the Truth-Value Judgment Task, Matsuoka (2005) emphasizes the importance of ethical consideration given to children, parents, and teachers (if data are collected at daycare centers or schools). As ethical issues are of growing concern in many research fields which may involve human subjects, research institutes throughout the world are establishing IRBs (internal review boards) to evaluate and approve research proposals. The “Ethics” section in the introduction of the current volume should be useful as researchers prepare applications for research grants or an IRB evaluation.

The limitation of the current volume is the reverse side of its strength discussed earlier. i.e. the approach that the editors chose: opting for the broad coverage of many different tasks, presented in a way accessible to researchers in different fields. The page limitation for each chapter no doubt made it difficult for authors to include a detailed discussion of the issues related to the tasks they covered, particularly when the chapter contained various experimental tasks. Since the content of each chapter is more or less an introduction to the rationale and procedures of the methods, readers who wish to learn about the validity of using those tasks to address their own interests will need to refer to the actual studies cited in the chapters.
4. Who Would Benefit from this Book?

As discussed in the previous section, the current volume is highly recommended to researchers who need to refer to or evaluate experimental studies. Instructors might find the current volume helpful in advising graduate/upper undergraduate students of psycholinguistics about experimental methods suitable for their research projects. As many of the methods covered in the current volume require expertise and intensive training as well as “hands-on” experience, however, researchers might want to use this book only to obtain a basic knowledge of experimental techniques and how they are used.

Overall, it is not only psycholinguists who can benefit from this book. By going through chapters of this book, readers can understand what type of data from what type of subjects can be obtained with current technology. It should provide “food for thought” for linguists as they consider the nature of the data used in linguistic analysis. Some methods, including Magnitude Estimation described in Chapter 3, can be used to supplement grammatical judgments of monolingual adults. In addition, the current volume includes an efficient review of central issues in comparative studies of monolinguals/bilinguals, typically-developing children and children with Specific Language Impairment, as well as deaf and hearing children. This should benefit linguists who would like to understand the nature of human language but do not have access to one or more of these subject groups. Finally, this book should be useful for readers who are generally interested in the study of second language acquisition to learn about different methods, as well as to see how methodological issues in L2 studies may be related to those in other subfields of linguistics.

5. Summary

The current volume provides an important and interesting reference for linguists interested in knowing which types of data can be obtained from current research methods. A basic knowledge of methodological issues, well summarized in the current volume, should help readers consider issues related to the design of experimental studies. Experimental data-collection cannot avoid the interference of language processing (cf. Pouscoulous et al. (2007)). The presentations in Part I will provide insights into how linguistic competence and language processing interact when either children or adults produce or comprehend language. Considering the likely interaction
between linguistic competence and language processing, experimental results from a variety of experimental methods can be used to supplement the use of grammatical judgments, which have long been a major (and controversial) source of data in generative grammar. The use of various experimental methods may ultimately provide fresh perspective on adult linguistic competence; such as Han et al.’s (2007) insightful claim about a possible “split” in adult grammar, based on experimental data. It seems, therefore, quite possible that the current volume will offer an important contribution to the advancement of theoretical linguistics. Moreover, Part II serves as a well-organized introduction to those readers not yet familiar with research projects which require comparisons between subject groups from different linguistic backgrounds. The compact and clear presentation(s) in each chapter should be of great help for researchers and students in various linguistic fields who have not yet received extensive training in experimental methods.

REFERENCES


[received December 20 2011, revised and accepted June 13 2012]

Faculty of Economics
Keio University
4–1–1 Hiyoshi Kohoku-ku
Yokohama 223–8521
e-mail: matsuoka@z7.keio.jp