

Analysing Learner Language 分析学习者语言

Rod Ellis Gary Barkhuizen

Analysing Learner Language 分析学习者语言

Rod Ellis Gary Barkhuizen



图书在版编目(CIP)数据

分析学习者语言/(新西兰)埃利斯(Ellis, R.)著.

一上海:上海外语教育出版社,2012

(牛津应用语言学从书)

ISBN 978-7-5446-2884-6

, I. ①分··· II. ①埃··· III. ①第二语言—语言学习—研究—英文 IV. ①H09

中国版本图书馆CIP数据核字 (2012) 第216681号

图字: 09-2011-691号

Every effort has been made to obtain permission, but in the event of negligence or particular difficulty of obtaining permisson we shall be pleased to come to a suitable arrangement with the rightful owner.

Analysing Learner Language was originally published in English in 2005. This bilingual edition is published by arrangement with Oxford University Press.

Licensed for sale in the People's Republic of China excluding Hong Kong, Macau and Taiwan.

本书由牛津大学出版社授权上海外语教育出版社有限公司出版。仅供在中华人民共和国境内(香港、澳门、台湾除外)销售。

出版发行: 上海外语教育出版社

(上海外国语大学内) 邮编: 200083

电 话: 021-65425300 (总机)

电子邮箱: bookinfo@sflep.com.cn

斌: http://www.sflep.com.cn http://www.sflep.com

责任编辑: 许进兴

印 刷:上海华业装璜印刷厂有限公司:

开 本: 890×1240 1/32 印张 13.125 字数 448 千字

版 次: 2012年12月第1版 2012年12月第1次印刷

印 数: 2000 册

书 号: ISBN 978-7-5446-2884-6 / H • 1404

定 价: 42.00 元

本版图书如有印装质量问题,可向本社调换

出版说明

本世纪初,外教社先后引进"牛津应用语言学丛书" (19种)和"牛津应用语言学丛书(续编)"(10种)。这 些图书由于内容权威、选择精当而受到了外语界的好评,在 科研论文中被广泛引用,对推动我国外语教学和研究的发展 起到了重大作用。

近年来,随着研究的不断扩展和深入,国内学界对研究资料有了新的需求,像"任务型教学法"、"英语作为国际通用语"、"二语习得的跨学科研究"等逐渐成为了热门的话题。有鉴于此,我们又从牛津大学出版社出版的应用语言学图书中精选了10本,以更好地满足广大教师和科研人员的需求。希望这次出版的这10本图书,能够和以前的29本一起,反映出国际应用语言学重要领域研究的前沿,为全面、深入推动我国外语科研起到新的作用,做出新的贡献。

Acknowledgements

The authors would like to thank Ute Knoch for her help in painstakingly checking bibliographical references.

The authors and publisher are grateful to the following for permission to reproduce extracts and figures from copyright material:

Athelstan for a screenshot from Barlow, M.: MonoConc Pro.

John Benjamins Publishing Company, Amsterdam/Philadelphia: 'A bird's eye view of computer learner corpus research' by Granger, S.; 'A corpusbased study of the L2 acquisition of the English verb system' by Housen, A. from Granger, S., Hung, J., and Petch Tyson, S.: Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching (p.5).

Blackwell Publishing: 'Negative feedback as regulation and second language learning in the zone of proximal development' by Aljaafreh, A. and Lantolf, J. P.: *Modern Language Journal* 78: 47.

Greenwood Publishing Group, Inc., Westport, CT: 'Discourse Completion Test'; 'Taxonomy of request strategies'; 'Frequency of use of request strategies' from Blum-Kulka, S., House, J., and Kasper, G.: *Cross-cultural Pragmatics: Requests and Apologies*. Copyright © 1989 by S. Blum Kulka, J. House, and G. Kasper.

Hodder Arnold (Arnold Journals): 'Metaphorical conceptualizations of ESL teaching and learning' by de Guerrero, M. C. M. and Villamil, O. S.: Language Teaching Research 6/2: 95–121 (2002).

Lawrence Erlbaum Associates, Inc: Table 3.2 from Ohta, A. S.: Second Language Acquisition Processes in the Classroom: Learning Japanese (2001, LEA) (p. 89).

Pearson Education (Longman): Fairclough, N.: Language and Power (p.110).

Penguin Books Ltd: Table 2a, 'How errors are corrected' from Tsui, A. B. M.: Penguin English Applied Linguistics: *Introducing Classroom Interaction* (Penguin English, 1995) (p.52). Copyright © Amy B. M. Tsui, 1995.

TESOL, Washington, DC: 'Creative construction in second language learning and teaching' from Burt, M. K. & Dulay, H. C. (eds.): On TESOL '75: New Directions in Second Language Learning, Teaching and Bilingual Education (pp. 21–32).

此为试读,需要完整PDF请访问: www.ertongbook.com

Université Catholique de Louvain, Louvain-la-Neuve, Centre for English Corpus Linguistics: Dagneaux, E., Denness, S., Granger, S., and Meunier, F.: *Error Tagging Manual Version 1.1* (1996).

Sources:

Yuan, F.: The effects of planning on language production in task-based language teaching.

Although every effort has been made to trace and contact copyright holders before publication, this has not been possible in some cases. We apologize for any apparent infringement of copyright and if notified, the publisher will be pleased to rectify any errors or omissions at the earliest opportunity.

Preface

This book has grown out the authors' experience of teaching SLA to post-graduate students in a wide variety of contexts—the UK, South Africa, Japan, the USA, and New Zealand. It was developed to replace traditional second language acquisition (SLA) courses based on lectures on key topics with a more hands-on, do-it-yourself approach, where students were required to grapple with second language (L2) data in order to address the kinds of research questions that have figured in SLA research. The materials in the book have now been used with a number of postgraduate students at the University of Auckland. We are grateful to these students for the valuable feedback they have provided.

Analysing Learner Language serves as an introduction to SLA research for postgraduate students and teachers wishing to undertake empirical studies of L2 acquisition. It has the following aims:

- 1 to familiarize readers with different methods for analysing learner language as expression and as content;
- 2 to examine the theoretical and research bases for the different methods; and
- 3 to develop readers' ability to undertake the analysis of samples of learner language using different methods.

The book provides full examples of the different methods for analysing learner language and also tasks for readers to practise the methods themselves. The chapters detailing the different methods of analysis (Chapters 3–13) have four main sections. The first provides an account of the historical and theoretical background of the method. This is intended to provide a rationale for the method and to demonstrate the contribution it has made to the study of L2 acquisition. The second offers a step-by-step account of the method together with an illustration of its application to actual data. The third section provides an example of a study that has employed the method; the study is summarized and subjected to critical review. Lastly, the fourth section is a task where readers are provided with a sample of learner language which they are invited to analyse using the method. The purpose of this section is to give hands-on experience in the actual analysis of learner language. Readers can try out each method for themselves.

However, Analysing Learner Language is not a research methodology book in the traditional sense of this term. That is, it does not address issues relating to the choice of a particular research methodology or of research design. Nor is it primarily concerned with methods of data collection. The thrust of the book is an account of the tools that researchers have used to analyse the spoken and written texts produced by L2 learners. The use of these tools in SLA has been motivated by the particular research questions that researchers have addressed. As the nature of these questions has changed over time, so new methods of data analysis have been developed. Thus, an account of the methods of analysis in SLA will necessarily involve consideration of the historical and theoretical contexts of each method. Thus, by learning about the tools for analysing learner language, readers will also be introduced to the key issues that have motivated SLA enquiry.

Because the book provides an overview of key areas in SLA research (by way of contextualizing the different methods of analysis), it can also serve as an introductory SLA text. It may appeal to teachers of SLA who wish to offer a more 'hands-on' approach to teaching SLA in place of the traditional exposition of the main areas of study in SLA. One way of learning about how learners acquire an L2 is by studying the language they produce. Finally, the book may also be used as a reference book for SLA researchers wanting to review the possibilities for data analysis before finalizing a research proposal.

Rod Ellis Gary Barkhuizen Department of Applied Language Studies and Linguistics University of Auckland

Contents

Ack	V	
Pre	i	
1	Introduction	
2	Collecting samples of learner language	1:
3	Error Analysis	5
4	Obligatory occasion analysis	73
5	Frequency analysis	93
6	Functional analysis	111
7	Analysing accuracy, complexity, and fluency	139
8	Interactional analysis	165
9	Conversation analysis	197
10	Sociocultural methods of analysis	229
11	Coding data qualitatively	253
12	Critical approaches to analysing learner language	277
13	Metaphor analysis	313
14	Computer-based analyses of learner language by Michael Barlow	335
15	Conclusion	359
Bibl	371	
Inde	397	

1 Introduction

The primary purpose of this book is to provide an account of what is known about how learners acquire a second language (L2) by introducing readers to the methods that have been used to analyse *learner language*. A secondary purpose is to equip readers to carry out analyses of learner language by themselves for research purposes.

There are a number of books that provide overviews of 'second language acquisition' (SLA) as a field of study (for example, Ellis 1985a and 1994; Gass and Selinker 1994; Larsen-Freeman and Long 1991; Mitchell and Myles 1998; Towell and Hawkins 1994). There are also books that review specific areas of SLA—for example, Tarone (1988) on variability in L2 learning, Skehan (1989) on individual differences, Kasper and Kellerman (1997) on second language learners' communication strategies, Gass (1997) on the role of input and interaction and Doughty and Williams (1998) on the role of form-focused instruction. Thus, the field is well-catered for in terms of published works documenting what is known about L2 learning and L2 learners.

As a field with close links to the social sciences, education and applied linguistics, SLA is also well-provided with general accounts of research methodology. In addition to established works such as Ary, Jacobs, and Razavieh (1990), Brown (1988), Cohen and Manion (1994), Neuman (1994) and Seliger and Shohamy (1989), there are a number of more specific books addressing the applications of research methodologies to SLA, such as Brown and Rodgers' (2002) survey of qualitative and quantitative research techniques, Faerch and Kasper's (1987) collection of articles on introspective methods, Brown's (2001) practical guide to the use of surveys in second-language related research and Gass and Mackey's (2000) account of the use of stimulated recall in SLA. Other works (for example, Hatch and Lazaraton 1991) offer explanations of the statistical procedures available to SLA researchers. Together, these books provide detailed accounts of the theoretical underpinnings of different research traditions, the main research approaches and the designs associated with them, methods of data collection and ways of analysing data.

Given the wealth of the published literature on SLA and on research methodology, it seems advisable to us to begin this book by explaining why we have taken the trouble to write this book. In *Analysing Learner*

Language we want to bring together a substantive account of SLA as a field of enquiry through an examination of the methods of data collection and analysis that have informed research in this field. That is, our main goal is to introduce readers to what has been discovered about L2 learning and L2 learners through a consideration of the methods that have produced this knowledge. Indeed, our assumption is that understanding SLA requires an understanding of its established findings (what Long (1990b) calls the 'facts' of SLA) in relation to the ways in which these findings have been established. Thus, we seek to introduce readers to SLA by familiarizing them with the various methods of data collection and analysis SLA researchers have employed.

This approach has been adopted in response to one of the authors' experiences in teaching SLA courses over a period of some fifteen years. The approach he adopted was initially a fairly traditional one, involving lectures on selected topics (error analysis, variability, language transfer, input and interaction, learning strategies, etc.) supported by a range of tasks that required some form of application of the material presented in the lectures. These tasks included a number of data analysis activities. For example, in one such task, students were asked to analyse a set of negative utterances produced by one L2 learner over a nine-month period. In another, they were asked to examine how teachers addressed linguistic form in the context of interactions derived from communicative activities. It was observed that these tasks frequently led to lively discussions of key issues in SLA and, crucially, to problems of interpreting what the data showed. In response to this, a decision was taken to redesign the SLA course around these tasks. This required a new way of structuring the content. Instead, of basing the course on topics in SLA, the new course was organized around the different ways in which learner language has been analysed over the thirty-plus years that SLA has existed as a recognized field of study. The purpose, however, was not so much to train students in how to carry out these methods of analysis (although this was, hopefully, a useful outcome of the course) as to involve the students in 'doing SLA' so that their understanding of its findings were grounded in a hands-on experience of how they were obtained. This book is a direct outcome and extension of this course.

This book, then, is an account of how SLA researchers have set about analysing learner language, of the theoretical positions that underlie their enquiries and of the main empirical findings that have resulted from them. In addition, indeed as an essential part of the whole enterprise, the book provides opportunities, in the form of data-based tasks, for readers to apply the different methods of analysis themselves. It is also important to note what the book does not seek to do. The book does not consider culture learning, although it does examine the role of cultural factors in language learning. (See, for example, Chapter 10.) It does not aim to provide advice to teachers about how they should teach. Obviously, though, an understanding

of how learners learn an L2 should inform how teachers teach and some chapters include tasks involving data taken from instructional contexts. Also, the book does not examine curricula for teaching an L2.

In this chapter, we briefly explain what we mean by 'SLA', which we will interpret very broadly. We will also consider what is meant by 'learner language', identifying two distinct senses of this term. Finally, we will attempt to demarcate the particular aspect of SLA research methodology we are concerned with (i.e. data analysis) by placing it within a broader framework of the nature of research and the different research paradigms that have been employed in SLA.

SLA

Two different senses of SLA need to be distinguished. The term is frequently used to refer to the learning of another language (second, third, foreign) after acquisition of one's mother tongue is complete. That is, it labels the object of enquiry. The term is also used to refer to the study of how people learn a second language; that is, it labels the field of enquiry itself. This dual use of the term is unfortunate as it can create confusion. In this book we will use 'L2 acquisition' as the label for the object of enquiry and 'SLA' as the label for the field of study.

SLA is multi-disciplinary. That is, it draws on insights and methods of research from a range of disciplines, including linguistics, sociology, sociolinguistics, psychology, psycholinguistics and education. This multidisiplinary aspect of SLA is widely accepted and is reflected in all the published surveys of the field referred to above. It has advantages and disadvantages. It affords a rich account of what is a highly complex phenomenon. But, as Gass and Selinker (1994) point out, it sometimes results in a failure of communication by scholars committed to different approaches. SLA is characterized by a host of controversies (see Block 2003 for a thoughtful analysis of some of the main ones), many of which are not resolvable as they derive from incommensurable theoretical positions. In this book we will seek an inter-disciplinary perspective. Indeed, such an approach is inevitable, given our intention to present a range of different approaches for analysing learner language, as the different approaches derive from different disciplines and different theoretical orientations. We will acknowledge controversies where they exist, but we will not attempt to resolve them. Nor will we stake out our own 'preferred' position, although, of course, we cannot promise to guard entirely against our own biases.

The study of how people learn a second language involves both an examination of those aspects of learning that are common to all learners the universals of L2 acquisition—and of those contextual and personal factors that explain the enormous variation in speed and ultimate level of attainment of different L2 learners—the individual differences in L2

learning. Again, the published surveys of SLA acknowledge both aspects. It is probably true to say, however, that they have devoted more attention to the universals of L2 acquisition than to social and individual differences. the latter aspect being generally consigned to one or two chapters after the former has been exhaustively treated. Ideally, these two aspects need to be integrated into a single theory of L2 acquisition. However, we are a long way from achieving this, although a number of recent publications on individual differences (for example, Onwuegbuzie, Bailey, and Daley 2000; Skehan 1998a; Robinson 2002) have attempted to discuss specific variables responsible for variation among learners in terms of a general model of L2 acquisition. In Analysing Learner Language we will not attempt an integrated theory but instead present methods of analysis that relate to both the establishment of the universal properties of L2 acquisition and to individual differences in L2 learning. All of these methods are capable of addressing both the universal and differential aspects of L2 acquisition. For example, the analysis of learner language in terms of fluency, complexity and accuracy (see Chapter 7) provides a means of identifying how external factors that shape the learning environment affect the language that learners produce and also a means of researching how individual learners' communicative styles differ. Nevertheless, it is the case that particular methods of analysis have been closely associated with either a nomothetic and universalist orientation or an idiographic and hermeneutic view of L2 learning. Thus, for example, obligatory occasion analysis (Chapter 4) and frequency analysis (Chapter 5) have served to provide evidence of a universal route of acquisition, while critical approaches (Chapter 12) and metaphor analysis (Chapter 13) have led to insights about the different approaches to L2 learning manifested by individual learners.

In this book, then we acknowledge the multidimensionality of SLA, as reflected in the variety of approaches to analysing learner language. We recognize, too, that SLA must include both an examination of the universal properties of L2 acquisition and an account of the social and personal factors responsible for individual differences among learners.

Learner language

Learner language is the oral or written language produced by learners. It serves as the primary data for the study of L2 acquisition, although, as we will see in Chapter 2, it is not the only type of data available to SLA researchers. We will also see that learner language is not a monolithic phenomenon but rather highly variable, raising important issues to do with what kind (or kinds) of learner language constitute the most valid data for

the study of how learners learn. These issues are of such theoretical importance, however, that we will briefly examine them here.

For many SLA researchers the goal of SLA is the description and explanation of L2 learners' competence and how this develops over time. Definitions of competence vary but all see it as involving underlying systems of linguistic knowledge (Canale and Swain 1980; Taylor 1988). We can ask, then, how the study of learner language can provide information about learners' underlying linguistic knowledge. To address this question we need to consider exactly what we mean by 'linguistic knowledge' and then to examine the relationship between competence and performance.

Linguistic knowledge is constituted as shown in Table 1.1. The central distinction is between implicit and explicit linguistic knowledge. Implicit knowledge is the kind of knowledge we possess of our mother tongue. That is, it is unconscious and is proceduralized so that it is available for automatic use in spontaneous production. Implicit knowledge consists of formulaic chunks (for example, 'I don't know' and 'How do you do?') and also unconscious knowledge of abstract patterns or rules relating to both

Type of knowledge	Sub-types			Definitions	
Implicit (procedural)	1	formulaic		Sequences of elements that are stored and accessed as ready-made chunks.	
	2	rule-based		Unconscious knowledge of major and minor schemas consisting of abstract linguistic categories realizable lexically in an indefinite number of sentences/ utterances.	
Explicit (declarative)	1	analysed		Conscious awareness of minor and major schemas.	
	2	metalingual		Lexical knowledge of technical and non-technical linguistic terminology.	

Table 1.1 Types of linguistic knowledge

minor schemas as in simple collocations (for example, V+verb complement, as in 'suggest + Ving' and 'demand + Vinfin') and to major schemas (for example, the construction of relative clauses). Explicit knowledge is conscious and declarative (i.e. it takes the form of encyclopaedic facts about a language). It consists of both conscious awareness of the same minor and major schemas that figure in implicit knowledge (but represented in an entirely different mental form) and of the metalanguage that can assist in verbalizing this analysed knowledge. (See Ellis 2004.) As we will see, learner language is capable of providing information about both types of knowledge but it is often difficult to decide which type of knowledge is reflected in learner production. While SLA researchers argue about the role that explicit knowledge plays in both acquisition and language use—compare, for example, Krashen (1994) and N. Ellis (2002)—they broadly agree that the linguistic knowledge comprising competence is essentially of the implicit kind and that the main goal of SLA is to account for learners' implicit knowledge.

A learner's implicit knowledge (competence) is not open to direct inspection. We cannot easily look into someone's mind to see how knowledge of language is represented or what kind of knowledge is being utilized in the performance of a language task. To some extent this is possible through magnetic resonance imaging. (See, for example, Chee, Tan, and Thiel 1999.) This technology enables us to identify those parts of the brain that are activated in performing a language task and, in the long run, may enable us to determine the parts of the brain responsible for different kinds of linguistic processing. However, we are a long way from being able to plot cortical organization with reference to language use in this way. Thus, by and large, researchers are forced to infer competence from some kind of performance. How learners perform some kind of language task serves as the principal source of information about what they know about the language.

This raises the general question about what kind of performance provides the most reliable and valid source of information. Here we find major differences in opinion. On the one hand some SLA researchers choose to rely on learner intuitions (in the form of judgements about the grammaticality of sentences presented to them) to discover what they know. Other researchers, especially those of a more functional orientation, prefer to collect samples of learner language. Not surprisingly, analyses based on grammaticality judgements and on learner language frequently produce different results. A learner may succeed in judging a sentence correctly as grammatical and yet be unable to produce the structure exemplified in the sentence in free production. Furthermore, learner language is itself not homogenous but rather highly variable, depending on both social factors (for example, whom a learner is addressing) and psycholinguistic factors (for example, the degree of attention that a learner is paying to correctness of form) and the

inter-relationship of the two. Linguistic performance, then, is inherently heterogeneous.

What solutions are there to this problem of variability? There are several:

- Redefine competence as itself variable. (See, for example, Tarone 1983 and Ellis 1985b.) That is, variability in performance is seen as reflecting a variable competence.
- Identify one type of performance as the preferred source of information about competence.
- Recognize the need for multiple sources of performance data and look for points of confluence as evidence of what a learner knows.

It is not our purpose in this book to enter the thorny debates that surround which of these three solutions to choose. The fact that we have elected to address how to analyse learner language (rather than data derived from some kind of test), however, reflects our belief about the centrality of this kind of data. In Chapter 2, we examine the different methods for collecting samples of learner language, pointing to what is now widely acknowledged in SLA, namely the need for data that reflects as closely as possible 'natural' language use (i.e. language that is situationally and interactionally authentic) while recognizing that the limitations facing the collection of such data often obligate researchers to resort to clinically elicited data (for example, by using pedagogic tasks). Here, though, we want to emphasize that there is no easy solution to the data problem and that the main requirements that should be placed on researchers are to specify explicitly what kind(s) of data have been collected and to justify the validity of these data in terms of a clearly stated theoretical position regarding the relationship between performance and competence (i.e. to address the validity of the data).

In the foregoing discussion we have implicitly treated learner language as expression but it can also be viewed as content. That is, we can view learner language in two entirely distinct ways. We can see it as providing evidence of what learners know about an L2 by examining the linguistic forms they produce. We can also view it as a set of propositions relating to whatever topics are being communicated about. These topics can include those relating to the second language itself—that is, learners can inform us about their beliefs and attitudes to the target language and to the target language community and about the behaviours they engage in when learning the language. This distinction between learner language as expression and as content is central to this book. In part, it relates to a distinction we have already mentioned—that between the branches of SLA that have focused on the universal properties of L2 acquisition (the psycholinguistic orientation) and those other branches that have addressed the factors responsible for individual differences in learning (the social or psychological orientation). By and large, learner language treated as expression has served as data for

investigating the universal properties of acquisition while learner language viewed as content has provided information about how learners differ in their attitudes and approaches to learning. Of course, as we have already pointed out, it would be possible to examine individual factors in terms of differences in the formal properties of the language produced by different learners and, similarly, it would be possible to build a picture of the universal aspects of L2 learning from learners' reports about learning. In general, though, this has not happened. Learner productions viewed as expression tell us what learners do with the language and thus do not readily shed direct light on such factors as social identity, learning styles, motivation, language aptitude or learning strategies. To investigate these, we need learners to self-report. Conversely, learners are unlikely to be able to tell us in what order/sequence they acquired grammatical structures (for example, whether they acquired unmarked/protypical forms before marked/prototypical ones). Nor will they have much idea why they acquire forms in the order/ sequence they do.

There is, therefore, a strong rational basis for distinguishing learner language as expression and as content. There are also empirical grounds. The methods of analysing learner language we will discuss in the following chapters divide quite clearly into those associated with expression and content. Thus, methods such as error analysis (Chapter 3) and interactional analysis (Chapter 8), borrowed from linguistics, view learner language as expression while other methods, taken from the social sciences, such as critical analytical approaches (Chapter 12) and metaphor analysis (Chapter 13), treat learner language primarily as content. Our claim is that we need to orientate to learner language in both ways to obtain a full account of L2 acquisition.

Research paradigms

The focus of this book is on 'data analysis'. We recognize, however, that data analysis does not occur in a vacuum but is an integral part of the research process. It is shaped by the purpose of the research and the theoretical principles that govern the chosen method of enquiry. In this section, we consider how the analysis of learner language fits into the broader research picture.

Table 1.2 outlines the key differences in the two research paradigms widely recognized in discussions of research methodology in the social sciences. These two paradigms have been variously labelled (for example, quantitative/qualitative; confirmatory/interpretative; positivist/non-positivist; nomothetic/idiographic; analytical-nomological/hermeneutic). The labels we have chosen are those of Cohen and Manion (1994)—normative/interpretative. However, following Neuman (1994), Table 1.2 also includes a third paradigm, critical research, which differs from the