ACQUISITION & LANGUAGE DISORDERS

EDITORS

Ilahsen

Lydia White

of Essex

McGill University

EDITORIAL BOARD
Anne Baker (University of Amsterdam)
Man (Max Planck Institut für Psycholinguistik, Nijmegen)
Katherine Demuth (Brown University)
Werner Deutsch (Universität Braunschweig)
Kenji Hata (UC Santa Cruz)
Hyams (University of California at Los Angeles)
Peter Jordens (Free University, Amsterdam)
Jürgen Meisel (Universität Hamburg)
Kim Plunkett (Oxford University)
Mabel Rice (University of Kansas)
Sheel Sharwood Smith (University of Utrecht)
Antonella Sorace (University of Edinburgh)
Karin Strumswold (Rutgers University)
Jürgen Weissenborn (Universität Potsdam)
Helmut Zobl (Carleton University, Ottawa)

Volume 18

© C. Klein and Gita Martohardjono (eds)

Development of Second Language Grammars
A generative approach

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA
Investigating Second Language Grammars

Some Conceptual and Methodological Issues in
Generative SLA Research

Elaine C. Klein                                      Gita Martohardjono
Queens College and the Graduate Center, City University of New York

Introduction

In May 1995 the Graduate Center of the City University of New York hosted GASLA95, the 3rd conference dedicated exclusively to research in Generative Second Language Acquisition, ten years after Suzanne Flynn and Wayne O'Neil inaugurated this research paradigm at MIT. Many of the papers presented at GASLA95 reflected the rapid changes and shifts in perspectives which were occurring at the time. On the one hand, there was a conceptual shift in generative second language acquisition (GSLA) which moved explanation away from an earlier dichotomy characterizing second language acquisition as either having or not having access to principles and parameters of Universal Grammar. In its place, more subtle and complex explanations were being formulated to account for development in L2 acquisition; these tended to investigate putative L1 influence on the initial state of the L2 learner’s mental representation of the target language. At the same time, the focus of research itself had also shifted to an investigation of functional categories which had by now been postulated as the locus of parametric differences. On the purely theoretical front important changes were also taking place, and Chomsky’s reformulation of GB theory into the Minimalist Program was inevitably influencing the direction of SLA research.

In an attempt to capture these paradigmatic changes we have collected a selected number of GASLA95 presentations into this volume. The contributions were thus chosen primarily on the basis of whether and how they reflected the
paradigmatic shifts occurring in the field. In this sense, the chapters in this volume together represent “state of the art” research in generative SLA. In addition, we have written this introductory chapter to raise further questions which we hope will contribute to the evolution of our field by providing new insights into both conceptual and methodological issues.

One of the issues which has driven much of GSLA research has been the search for a principled explanation of how non-primary language acquisition differs from primary language acquisition. In attempting to characterize these differences, GSLA research has largely concentrated on investigating differential representations of knowledge in native and non-native systems. This approach entails two viewpoints which are not always made explicit. The first is that representation of knowledge, i.e. of core grammar, constitutes the locus of fundamental cognitive differences between native and non-native language acquisition. The second is that the representation of core grammar constitutes the most significant aspect of the process of language acquisition itself. We will refer to this as the “differential representation approach to explanation in SLA.”

In this introductory chapter, we give a critical appraisal of differential representation approaches by pointing to some of their conceptual and methodological limitations. We will suggest adopting parallel research strategies which rather than concentrating on differences in representations of the target grammar focus on issues pertaining to the process of grammar construction and, more specifically, on the necessary steps involved in mapping from a non-instantiated UG to language-particular grammars. We argue that from a generative perspective, such research strategies are crucial in furthering our understanding of the nature of SLA in general, of development in SLA in particular, and of SLA/FLA differences. In addition, we offer some suggestions on how we can strengthen existing methodologies used in GSLA research to reduce the problems which are inherent in representation approaches, in particular measuring knowledge through language use.

We begin, in Section 1, with a brief summary of three major research questions in the field, which adopt the differential representation approach. Section 2 provides a rationale for shifting the focus of research from a definition of L2 knowledge and representation to a specification of process and mechanism in grammar construction. Here we also propose potential areas of investigation. In Section 3, we take up questions of methodology which need to be considered in GSLA research and give a classification of the methods used in the chapters themselves in this volume. We end, in Section 4, with a brief overview of the chapters.

1. Representation Approaches in Generative SLA Research

Even a cursory look at generative SLA research reveals that since the inception of the field, research has largely focused on differences in L1/L2 representation of grammar to explain divergences in native and non-native language use. This is particularly evident when we consider the major questions underlying GSLA research:

1. Do UG principles constrain SLA?
2. Is parameter resetting possible in L2?
3. What is the initial state of L2 acquisition?

Essentially, since for L1 acquisition it is assumed that 1. UG principles constrain learners’ hypotheses; 2. parameter-setting is possible, and 3. the initial state is a “pure”, i.e. uninstantiated, version of UG principles, these questions identify potential areas of divergence from natural language acquisition in the case of non-primary acquisition. In this section we will illustrate how all these questions focus on differences in grammar representation, while largely leaving out of the inquiry the mechanisms of the processes by which these representations are constructed. In Section 2 we argue that an exclusive concern with differential representation across native and non-native systems is based on an incomplete characterization of the acquisition process. Furthermore, we give specific examples of how a research strategy investigating in detail the processes by which non-native systems must arrive at target representations might give interesting insights into why native and non-native systems tend to be different in the first place.

1.1 Questioning the role of UG in non-primary language acquisition

We begin with a discussion of the earliest question in the field, that of UG accessibility. Although attention to this question has receded in recent years, its significance lies in proposing the most radical version of the differential representation approach to SLA.

Under the so-called “no-access” position (for a discussion of this position, see e.g. White 1989; Eubank 1991; Epstein, Flynn and Martohardjono 1996, amongst others) adult non-primary language acquisition is essentially viewed as an entirely different cognitive activity from that of child L1 acquisition. A principled explanation of L1/L2 differences is offered by postulating the
inaccessibility in adult L2 acquisition, of those cognitive principles underlying the construction of L1 grammar representation, i.e. Universal Grammar (see Gregg 1996 and Epstein et al. 1996 for a classification of the various positions on UG in SLA). This is a very strong position, since if one takes seriously the claims of UG theory — that a domain-specific cognitive module governs the human ability to acquire natural languages — the no-access position essentially claims that L2 acquisition does not fall within the domain of natural language acquisition.

In view of this, it is interesting to ask why the no-access position captured the imagination of researchers in our field which was, after all, primarily concerned with characterizing SLA from a UG perspective. Notice, for example, that in questioning UG accessibility, early SLA research stood in stark contrast to UG-based First Language Acquisition research. Here, the initial state in the child's grammar was assumed to be UG and divergences from adult grammar and use were explained in terms of various processes necessary to complete acquisition. For example, research focused on problems of learnability (e.g. Waxler and Manzini 1987), maturation (Borer and Waxler 1987), performance constraints (Fodor and Crain 1990; Valian 1990), and mapping from core grammar to language-specific grammars (e.g. Borer et al. 1995). With the possible exception of maturation, which is unlikely to occur in adult L2 learners, this same approach could also have been taken in UG-based SLA research. That is, as an a priori research strategy, we could have adopted the view that SLA, like FLA, is yet another instance of natural language acquisition, hence that it is by hypothesis UG-constrained, and then gone on to investigate why L2 learners take such a tortuous path to arrive at the steady state.

The view that L2 acquisition is UG-constrained constitutes the null hypothesis and this has sometimes been interpreted to mean that both the course and the result of native and non-native language acquisition should be identical. This position is of course immediately falsified, given the high occurrence of incomplete acquisition in SLA. But this strong entailment is in fact not warranted by the null hypothesis. All it predicts is that like native acquisition, non-native acquisition potentially converges on the core grammar representation required by the L2. Divergences from L1 acquisition — as evidenced, for example in development, use and attainment — could then be investigated from the perspective of differences in the corollary cognitive processes interacting with UG knowledge in the instantiations of the non-primary core grammar.

The fact that many researchers in the field did not automatically adopt the

null hypothesis reflects the field's general concern with a question which predates SLA research: a specification of the fundamental cognitive differences underlying child and adult language acquisition. This idea finds most developed articulation in Lenneberg's Critical Period Hypothesis (henceforth CPH; see for example Krashen, Scarr-Seppa and Long 1982; Long 1990). No-access positions in effect provided a principled explanation of purported CP effects in UG-theoretic terms and therefore needed to be taken seriously. In addition, two phenomena in SLA lent considerable intuitive appeal to no-access positions, namely the already mentioned and oft-noted failure by L2 learners to achieve native-like proficiency and systematic divergences from native speech at the near-native level. We will briefly discuss each of these in turn.

1.1.1 Non-attainment of proficiency

An undisputed difference between child and adult language acquisition is captured in the anecdotal yet compelling observation that adult second language learners often fail to attain native-like proficiency (variously referred to as "lack of ultimate attainment", "fossilization", "non-completion" and "lack of success", e.g. Schachter 1990; Bley-Vroman 1990). Of course, many researchers have reminded us that non-attainment in SLA has yet to be established as the norm, especially in the systemic domain, and that proficiency in more than one language is actually not that uncommon in multilingual and bilingual settings (see e.g. Wode 1991; Cook 1995; Flynn and Samuel 1991; Bielystok and Hakuta 1999). Interestingly, however, when post-pubertal learners evidencing full proficiency in the L2 are documented, this is often ascribed to neurological exceptionality (see e.g. Schmidtsman and Dorian 1988; Ohler 1989; Smith and Tonpi 1991). While the issue remains unresolved, "lack of success" and the related question of "variability in success" in SLA continue to be regarded as powerful arguments for cognitive differences between child and adult language acquisition and therefore for the CPH (see Long 1990).

1.1.2 Near-native studies

At the same time, studies on successful SLA learners, i.e. near-native speakers, also provide evidence for proponents of the CPH. From a cognitive perspective, systemic behavioral differences between native and near-native speakers constitute an inherently interesting area of research: By shedding light on what is (non-) attainable in the limit for the non-native speaker for whom we can assume L2 steady-state attainment, such studies potentially point us in the
direction of precisely those aspects of the acquisition process where L1 acquisition critically differs from adult L2 acquisition. It should be noted that while systemic differences have been isolated, they are said to occur primarily in non-core aspects of grammar (e.g. Coppingers 1987; Searce 1993; Ioup 1994; Birdsong 1992).

Nonetheless, taken together, the observed rarity of native-like proficiency and the systematic divergences of use at the near-native level indeed provided compelling arguments for significant differences between L1 and L2 acquisition. By tracing the roots of such differences to fundamental cognitive differences, and in particular to an inability to access and use the domain-specific module needed for grammar construction in SLA, no-access positions articulated the most radical (and arguably the most interesting) differential representation approach (e.g. Bley-Vroman 1990; Clahsen and Mayrken 1986; Meisel 1991).

Often, however, the most radical positions are also the most difficult to defend and, ultimately, proponents of no-access positions must show that differences in L1/L2 acquisition and differences in native/non-native use of the L2 occur in areas of grammar which derive directly and uniquely from UG. This would strengthen their claim that such effects necessarily stem from the absence of domain-specific knowledge, and not, as is equally plausible (and arguably more probable), from other, equally critical factors in the acquisition process. In this respect, we hope to show (Section 2) that L2 acquisition effects are just as likely to be the result of corollary processes driving the mechanism of grammar construction and, consequently, that the investigation of such processes is critical in GSLA. We are, of course, not the first ones to make this observation. Almost a decade ago White (1989) pointed to this possibility, as did subsequently Schwartz (1991), Flynn and Manuel (1991) and Martohardjono and Flynn (1994). Nonetheless, relatively little research on such issues has been conducted in the field of GSLA to date (cf. Fernández this volume, Flynn et al. 1998; Harrington and Juffs 1996; White and Juffs 1998).

1.2 Parameter-resetting and Initial state studies

As evidenced by the chapters in this volume, current GSLA research has largely shifted away from the access/no-access debate to the question of whether in spite of UG access parameter-setting is possible in SLA, and to investigations of the L2 initial state as a source of explanation of divergent L2 patterns. On the one hand, this type of research makes qualitatively different hypotheses from no-
and Young-Scholten 1994, 1996; transfer of lexical and functional categories without feature specification: Eibank 1996 among others), and adoption of the L1 grammar in all and only those cases where it can provide a pattern for the target grammar (White 1995). Notice that again the source of the explanation for divergence in L2 development is hypothesized to be the learner's extant representation of the L1 core grammar which is presumably different from that of the L2 and hence problematic. In summary, even as they view L1 and L2 acquisition as deriving from the same cognitive module and being constrained by UG, parameter-resetting and initial state studies continue the practice established by no-access positions, of focusing on differences in L1 and L2 knowledge and representation to explain L2 acquisition.

Perhaps one of the reasons differential representation approaches have been so readily embraced in GSLA research, is that the generative paradigm which our field is based on revolves primarily around the specification of knowledge and representation of language. It has in fact been argued that representation approaches provide more principled explanations of development in SLA (e.g. Schwartz 1995; Clahsen and Mayr 1995) as they do for development in FLA (Hyams and Waxler 1994). Nonetheless, there are both conceptual and methodological limitations to a strictly representational approach when trying to solve the acquisition puzzle. Close inspection of a UG-based acquisition model provides compelling reasons to extend our research paradigm to include a more precise investigation of the steps involved in grammar construction, i.e. when learners map from UG to language-specific grammar, as well as questions of new performance factors, such as parsing, in context with knowledge of principles in the restructuring of intergrammar. Furthermore, since differential representation approaches necessarily rest on the competence/performance distinction, they demand more refined research methodologies from the ones that are largely in use. In sections 2 and 3 we will treat each of these questions in turn.

2. A UG-based acquisition model

What motivations are there to look beyond differences in the existing representations of native and non-native grammars to explain SLA from a generative perspective? In order to answer this question, it might help to review a general UG-based model of language acquisition (that is, not restricted to SLA). Broadly speaking, such a model must be based on the following premises:

1. Principles and parameters of UG constitute the domain-specific knowledge which guides human language acquisition by constraining the representation of language-particular core grammars.
2. However, the language acquisition process cannot be determined by knowledge of UG alone. Rather, in the specification of language-particular core grammars, UG knowledge must interact with a learning mechanism to allow convergence on a particular representation of grammar (i.e. the one required by the target language).
3. Input from the ambient language is critical to allow convergence on the right (i.e. target) representation.
4. The grammar construction mechanism must encompass a language parser to process the input. In addition, in order for the mechanism to operate properly, certain conditions must be met to satisfy (some specification of) learnability theory, i.e. a theory of how the input interacts with UG knowledge to allow convergence.
5. Language acquisition can be characterized as a sequence of restructuring grammars and the developmental changes observed during the course of acquisition are the result of grammar restructuring.
6. Restructuring is driven by input. In particular, during the process of language-particular core grammar specification restructuring only takes place if the input is not compatible with the grammar that is available.
7. Completion is attainment of the steady state. At this point, no more restructuring occurs.

The model is schematized as follows:
2.1 Applying a UG-based acquisition model to SLA: Property vs. process

The question we are now posing is the following: Is it more important to look at the properties of L2 grammars or the process by which these grammars are constructed in the acquisition process? From the perspective of no-access positions, this is of course a moot point. Since here, by hypothesis, SLA does not derive from the knowledge or mechanisms driving FLA, supporters of the no-access position would simply not investigate SLA from the perspective of a model such as the one outlined above, so that the question of whether it is more important to focus on property or process issues is peripheral at best.

If, on the other hand, one takes the position that SLA follows the same principles as FLA, in that it is subject to constraints imposed by UG (the position adopted by parameter-setting and initial state studies), then the distinction between property and process issues becomes central. In particular, the question now is the following: If UG is accessible to constrain grammar representation in SLA, then what accounts for the differences between primary and non-primary language acquisition (i.e. child L1 vs. adult L2), and what causes the differences between native and non-native use (adult L1 vs. adult L2)? Here an investigation of process issues could provide non-trivial answers. Suppose, for example, that failure to attain the target representation is indeed the case for SLA (i.e. the L2 steady state is not the same as the steady state of the L1 speaker). This would be a case of the learning mechanism failing to converge on the required grammar, so the relevant question to ask is what in SLA is different from FLA to make it so. What is important to evaluate here, in other words, is not that FLA is different from SLA, but in what ways the conditions required for the mechanism of grammar construction to operate properly are not met in SLA, clearly a question concerning process, rather than property.

Consider furthermore the question of L1 transfer asked in parameter setting and initial state studies: In trying to solve the problem of how L2 learners move from deviant or ‘less target-like’ behavior to ‘more target-like’ behavior, such studies concentrate on property issues by asking: What is the knowledge state of the non-native speaker at time x and is it the same as that of the native speaker? If not, how is it different? Is it the knowledge state required for the L1? Does the L2 learner eventually converge on the required knowledge state? Undeniably, these questions are important, yet their answers again critically involve an understanding of the process by which the learner moves from one knowledge state to the next. To illustrate, there seems to be
considerable evidence from parameter-setting studies, that focusing on L1 representation alone gives an incomplete picture of non-native systems. Most studies, for example, show that the predicted transfer of clusters of properties associated with a given L1 parameter setting does not occur in SLA (e.g., White 1985; Mendel this volume). Consider now how an investigation into process issues, in this case L2 parsing, might shed some light on this: Typically, a cluster of properties results in surface strings each with a different processing load. Such differences in turn might explain why the L2 acquisition of a set of structures deriving from the same parameter is not simultaneous in real time.

An accurate characterization of L2 development cannot only be concerned with capturing differing underlying knowledge states, but must in addition investigate the factors which propel the changes in intermediate grammar representation. By focusing exclusively on purported differences in property issues, representation approaches risk overlooking what could be the most significant factors accounting for differences in L1 and L2 development. On the other hand, extending our investigations to process issues such as learnability, parsing, quality of the input and what would constitute triggers for restructuring (see Fodor, this volume), might well provide us with the key to understanding why knowledge states differ in L1 and L2 in the first place.

2.2. Research Questions in Investigations of Process

Let us now consider some research areas which are relevant to an understanding of process. Under the model outlined above, a change in grammar representation (i.e. restructuring) only takes place if the learning procedure is able to propel acquisition from one state of knowledge to the next. Furthermore, acquisition is error driven; that is, it proceeds only if input is incompatible with the available grammar. Thus, an examination of both quality and processing of the input are important areas of investigation for GSLA research. Take, for example, the parameter governing overt syntactic movement: Only after relevant input from the target language has been processed by the learning mechanism, can the entry for the functional category CP in the lexicon be specified as [+strong]. Overtly, certain sentence-types which are readily available in the input are sufficient to trigger restructuring. In the L2 acquisition of English, for example, wh-fronted object questions should be sufficient to trigger the [+strong] feature.

derivational analyses, including one which does not involve surface movement. Thus, other input strings are needed to trigger the movement analysis (see Martohardjono and Cair 1993; for a general discussion of problems posed by surface triggers, see Fodor, this volume). This suggests that an investigation of what constitutes necessary triggers in the L2 input is essential in GSLA research.

Differences in processing are also likely to exist between native speakers and L2 learners. Could such differences, rather than differences in representation, account for non-native behavior? Preliminary research by Flynn (1995), Juffs and Harrington (1995) and Fernandez (this volume) suggest this possibility. In this regard, a comparison of processing between child L1 learners and adult L2 learners becomes essential. It is possible, for example, that input in L2 acquisition is not processed in exactly the same way as in L1 acquisition, which could in turn result in a failure for L2 input to become intact. This would mean that the necessary triggers to propel the L2 learner from one knowledge state to another are missing.

These are all areas which warrant serious consideration and investigation if we are to gain deeper insight into the nature of SLA, and in particular of SLA development. While the suggestions we have made above are preliminary in nature, and therefore rather sketchy, we hope to have shown that investigations of process issues constitute an integral part of GSLA research which need to be pursued actively and systematically.

3. Methodological Issues in Generative Second Language Acquisition Research

So far we have argued that representation approaches are limited because of their narrow focus on knowledge states to explain development and use. We argued on conceptual grounds for a more comprehensive investigation of the process of acquisition. Also crucial to this investigation are more rigorous standards in the methodologies we employ. We turn to this issue next.

Perhaps the biggest problem facing GSLA research is having to evaluate competence through performance (see Cairns, this volume, for similar problems in FIA research). The fact that all performance phenomena, whether instantiated as judgments in metalinguistic tasks, or as elicited or natural speech production, are necessarily influenced by extragrammatical factors (see Martohardjono 1998i
least minimizing, the effects of such extraneous factors.

It is reasonable to expect higher degrees of performance effects in learners than in native speakers and we have argued elsewhere that in measuring learners' adherence to UG principles, we should not expect statistical non-significance in the way for example learners reject ungrammatical sentences, compared to native speakers. Instead, measuring multiple sites in the L2 grammar and evaluating its internal systematicity as a UG-constrained grammar could be used as a gauge of UG-adherence (see Martohardjono 1999; Li 1998; Dekydtspotter and Sproat 1996). This research strategy is similar to the one adopted in studies examining cluster effects in parameter resetting (e.g. Montrul, Prevost, this volume).

Another way to control for performance effects is to use multiple tasks in controlled experiments. Performance effects are likely to be different for different tasks, whereas knowledge should remain constant across tasks. By using a multiplicity of tasks on both L2 learners and native speakers, it might be possible to tease apart knowledge from task effects. In this section we consider, in detail, ways to refine and improve research designs so that they more accurately and reliably measure competence through performance.

Although researchers may claim that findings from L2 studies contribute answers to particular theoretical questions, it is often the case that such findings are artifacts of a particular experimental design, procedure, test or stimulus type. In our efforts toward refining L2 methodology, central questions arise when we review studies such as those reported in this volume. Among some questions to be addressed in this section are the following:

1. To what extent does performance in an experiment tell us, among other things, about a learner's:
   a. Developmental processes and/or steady-state grammar?
   b. Knowledge of language versus use of language?
   c. Knowledge of grammatical principles versus processing strategies?

2. In the analysis and interpretation of empirical data, to what extent does the test design promote consideration of the following:
   a. Optional as well as variable learner choices?
   b. Preferences versus categorical learner choices?
   c. Measures to encourage test reliability?

These questions shall be briefly discussed in turn.

3.1 Investigating developmental processes and/or the steady-state grammar

In experiments which attempt to test the steady-state grammar, selected participants must of course show evidence of having completed the language acquisition process, i.e. stabilized their grammar. The obvious difficulty in making such determinations is one of the reasons many recent studies concentrate on the investigation of L2 developmental processes, as represented by most of the studies in this volume.

Cross-sectional studies like those of Montrul and Roberts and Sonce (this volume) attempt to investigate the stages of acquisition by testing a variety of proficiency levels. Clearly, an objective of many studies is to capture learners at the first stages of a particular developmental progression and then to illustrate the subsequent points in the process.

The difficulty of capturing the earliest intergrammatical stage or, as in many UG studies, the initial setting of a parameter, has led to many inconsistencies in the so-called "initial-state" studies (for a critical review, see Epstein, Flynn and Martohardjono 1996a). Since a cross-sectional study requires that the same test be given to learners at the earliest as well as the most advanced intergrammatical stages, the problem of formulating a test so that it is easy enough for very beginning learners and difficult enough to reflect different stages is obviously burdensome. Multiple tasks are a requirement for such research as mentioned above; optimally, a longitudinal study can also reinforce interpretation of the data.

Many of the studies in this volume test knowledge of structures that only begin to appear at later stages in the intergrammatical, studies which are not especially concerned with capturing learners' earliest grammatical hypotheses (e.g. In this volume see Fernández who investigates relative clauses; Pérez-Leroux and Li who investigate relative clauses and noun complements). In Fernández' study, for example, learners at "high intermediate" proficiency levels were tested, obviously not reflecting their (potential) end-state grammars. Such research promotes further study of the possibility of subsequent developmental stages. By contrast, Pérez-Leroux and Li test both "intermediate" and "advanced" L2 learners to illuminate developmental differences which come at the later stages of acquisition. Optimally, test design should ensure that enough learners are tested at each stage so that group results can be meaningful. The importance of analyzing individual results, as well, will be discussed below.
3.2 Distinguishing knowledge of language from use of language

As already noted, grammatical knowledge is different from grammar use, a distinction which becomes blurred in research which must rely on learner performance. Such performance is tested through production and perception tasks, the first of which reveals the learner’s output either orally or in writing.

Production tasks such as written sentence completions and translations (e.g., White et al., Chapter 7) and oral spontaneous conversations (e.g., Liceras et al., Chapter 5), clearly provide an added burden to the learner; that of using grammatical knowledge for the purposes of producing a test answer, communicating a message, etc. In addition to the usual confounding factors which may result in variation in performance for a single learner (e.g., fatigue, slips of the tongue, emotional states), there are others: Production tasks often depend, for example, on the degree to which learners: a) exhibit certain temporal patterns, e.g., risk-takers versus perfectionists; b) command such linguistic devices as articulatory precision and writing ability, and c) handle cognitive and processing complexities (with different tasks varying in cognitive and processing load — see Chaudron 1985).

Considering these factors, it is important to analyze how native speaker controls perform across tasks, determining the degree to which their performances vary and comparing this variation with that of non-native speakers. It is not surprising, however, that L2 learners’ output might not only be variable within and across production tasks, but that such performance falls far short of their underlying grammars.

Controlled production tasks, nevertheless, serve an important purpose and are more useful for the researcher than communicative tasks. Spontaneous production is often hindered by well-known avoidance factors and involves communication which is governed by entirely different principles. On the other hand, elicited production tasks offer a view into learner grammars by informing us which structures these grammars can actually generate; as such, production tasks must supplement perception tasks which tell us the structures that these grammars can process and interpret. Thus it is very important, from a pedagogical standpoint and otherwise, for researchers to uncover discrepancies between learner grammars and learner performance using both kinds of tasks, to find patterns that exist and seek non-trivial explanations for these gaps. Exactly how this may be empirically investigated is a major challenge for L2 research designs.

Minimally, production tasks should go hand-in-hand with perception tasks as some studies here illustrate (e.g., Prévost, Chapter 3; White et al., Chapter 7; Robertson and Sonee, Chapter 13). With fewer behavioral factors intervening, perception tasks have been argued to come closer to revealing a learner’s underlying knowledge.

To further this goal, some L2 studies reported here have followed native language research in their interpretation tasks (Ferrandez, Chapter 9, for example, uses written sentence interpretations based directly on adult language research). Many studies here and elsewhere test learners’ grammaticality judgments (e.g., Prévost, Chapter 3; Pérez-Loccic and Li, Chapter 6) which many have acknowledged to be problematic in formulation, administration and interpretation (see discussions, for example, inBirdsong 1989 and Ellis 1991).

Recently, efforts have been made to further improve such perception tasks. To refine their analysis of linguistic acceptability, Robertson and Sonee (Chapter 13) add the “numerical magnitude estimation” procedure to a contextualized grammaticality judgment task. To reveal differences between native and non-native speakers who superficially appear to have similar grammars, other L2 researchers have followed psycholinguistic experiments by measuring learner reaction times (e.g., Eubank 1993; Clahsen and Hoog 1993). While vague and conflicting results suggest that this methodology needs further refinement, researchers such as White and Juffs (1998) interpret reaction time results, among other findings, to suggest processing rather than grammatical differences between native and near-native speakers. It is this distinction to which we now turn.

3.3 Distinguishing knowledge of grammatical principles from processing strategies

In discussing grammatical knowledge in the broad view, the previous section further distinguished knowledge provided by universal principles from that which comes from learning and processing principles, and the interaction of the three.

Until recently, the domain of native speaker research alone, processing questions have begun to appear in L2 acquisition studies involving perception tests. Off-line tasks, which cannot differentiate between the grammar and parsing routines, are also being supplemented by on-line tasks which assume principle-based parsing (Prichett 1992). Measuring accuracy along with reaction times, such experiments attempt to identify the loci of learner difficulty; this is done by presenting stimuli sentences word-by-word and recording response latencies as
crucial sites (e.g., Juffs and Harrington 1995, 1996), adding further refinements to our knowledge of L2 performance.

The crucial question remains as to how such processing performance is related to learner grammars. To an increasing number of psycholinguists, the effects are direct: For example, Fodor (Chapter 14) outlines the contributions of the parsing mechanism to parameter-setting, arguing that the parser underlies grammar construction in acquisition. On the other hand, L2 researchers such as Chae and Hong (1995) appear to look at the issue from the opposite direction; they claim that processing studies can provide a direct view into a learner’s internal grammar, such that processing effects reflect the outcome of the grammar (rather than drive the grammar). Klein (Chapter 8) provides an overview of L2 processing studies and theories, and the developments in research which has led up to them.

3.4 Optionality and variability

In L2 research up to now there has been much discussion of learner variability (see, for example, Ellis 1990; Gregg 1990; Tarone 1983, 1990). Following the linguistic literature (e.g., Fukui 1995), the important issue of optionality has more recently been given attention in L1 and L2 studies (e.g., L1: Wesler 1994; L2: Eubank 1993a, Robertson and Sorace, Chapter 13 and references therein). If optionality and variability are inherent parts of an L2 learner’s grammar, experimental tasks must tap all L2 learner options as well as those of native speaker controls. It will also be crucial to distinguish between preferences and possibilities — see below.

To clarify our distinction between variation and optionality in a fairly simple way, we examine two types of hypothetical stimuli in a test of L2 English, first the ungrammatical examples shown in (1) and (2):13

(1) He looked the house.
(2) He looked the girl.

First and foremost, acceptance of an ungrammatical sentence, such as (1) and/or (2), does not necessarily mean the absence of a rule, as many L2 researchers have assumed; rather such acceptance can reflect a learner’s variable optional rule system.

In a grammaticality judgment task, we would interpret a pattern of learner acceptance of sentences like (1) but rejection of sentences like (2), or vice versa, to represent systematic variability. In our understanding, a variable rule is one that operates in some environments or under some conditions and not others.

We now add the grammatical examples shown in (3) and (4):

(3) He looked at the house.
(4) He looked at the girl.

Optionality would be evident if the learner pattern shows acceptance of (1) along with acceptance of (3), and/or acceptance of (2) and (4). We therefore interpret an optional rule to be one that sometimes does and sometimes does not operate in the same environment. Both systematic variability and optionality as described above can be said to be within the domain of performance or it can be argued that both of them reflect the learner’s competence. This is a different view from the generally accepted assumption that variability is solely a performance phenomenon while optionality suggests competing representations within the same learner.

Evidence of both variability and optionality provides crucial information about the L2 learner’s interlanguage stages and developmental process without necessarily measuring these against the target grammar. Therefore, we argue that L2 experiments should include stimuli that will test for these kinds of learner behavior, stimuli which should include grammatical sentences and their ungrammatical counterparts (as in (1) versus (3)). The additional inclusion of sentences which vary only in lexicon (as in (1) versus (2)) will not only tell us more about variation, within and across learners, but will also help us understand the interaction between structure and feature specification in the lexicon, crucial areas of research.

In addition, we suggest the following:

1. That there be repetition of lexical items whose features are predicted to interact with structure; that is, to ensure that the learner is responding to structure rather than to lexical differences, researchers must carefully control for lexical variation within sentence types or across types that will be compared. For example, Klein (1993) tested for acceptance of “null prepositions” in declaratives, questions and relative clauses using stimulus sentences with a repetition of lexical items, as in:

(a) The two friends talked the interesting movie last night.
(b) Which movie did the two friends talk last night?
(c) This is the movie that the two friends talked last night.

(From Klein 1993:71)
2. That individual results always be analyzed, since group analyses mask variability and optimality. Individual patterns of response can provide us with significant insights that are obscured when only group results are analyzed (see Cairns this volume for other implications of individual vs. group analyses). Insights into L2 learners' variable and optional rules, as noted above, along with evidence for varied paths of intergrammatic development, provide crucial acquisition information which can only be determined by comparing individual learner differences with group patterns. (See also footnote 15 which shows how studies report different findings when their methods of analyses are not comparable.) Both individual and group analyses are performed in most of the studies in this volume (see chart at the end of this section).

3.5 Categorical choices versus preferences

Cairns also cites the importance of distinguishing between grammars, which are about possibilities, and preferences which are the domain of performance theory. With examples from L1 studies on the acquisition of Control structures and Binding Principles, she argues that our testing procedures must enable us "to learn all possible interpretations of a given sentence for each subject."

The interpretation tasks in L2 studies of Binding Principles are notorious for falling in this regard. In their early study of reflexive snaphers, for example, Broselow and Finer (1991) tested learners on sentences like the following:

(6) Mr. Fat \{expects\} Mr. Thin to paint himself.

(7) Mr. Fat \{believes\} that Mr. Thin will paint himself.

(From Broselow and Finer 1991:51)

The researchers gave participants four pictures from which to choose a match for the meaning of sentences like those above; their intention was to discover whether the antecedent for the reflexive (himself) would be bound locally (to Mr. Thin) or long-distance (to Mr. Fat). Since respondents were not asked which picture/sentence matches were disallowed, their choices may simply have been preferences.\(^{15}\) Others have pointed to this problem: Lakshmanan and Terriishi (1994) criticize the binding studies of Thomas (1989) and Himkawa (1990), both of which offer multiple-choices for learner responses, a task which Lakshmanan and Terriishi argue "may be tapping only subjects' preferences when interpreting reflexives and not their syntactic judgments." (1992) Thomas (1991:231, fn. 37), improving upon her earlier study, correctly notes that until learners reject a binding pattern (e.g. non-local binding for English), we simply don't know whether responses are preferences or whether they reflect the underlying grammar.

Grammaticality judgments testing a variety of linguistic structures are not exempt from such concerns. A learner who rejects a sentence, or sentence type, has not necessarily made a categorical choice. The importance of analyzing individual results becomes transparently clear when the researcher finds that the same subject also accepts the same sentence or sentence type within a task or across tasks, suggesting optimality as discussed above. If the language itself permits alternative choices (for example, as in English L2 studies of pied-piping and preposition-ellipsis), and learners accept both choices but to differing degrees, preferences may be statistically measured for interesting and potentially important comparisons with native speaker norms.

Thus, in some cases, researchers may acknowledge their interest in preferences rather than categorical choices, but this should be explicitly stated. In Fernández's study (chapter 9), for example, L2 learners and bilinguals had to "circle the word that correctly answered each question, based on the meaning of the sentence directly above it." Choices like the following were presented:

\[
\text{8) Roxanne read the review of the play that was written by Diane's friend.}
\]

What was written by Diane's friend? \{the review the play\}

Fernández interprets her results in terms of "preferred" responses since both answers are permitted by the Spanish L1 and the English L2. However, such preferences appear to differ among native speakers of the two languages and the researcher focuses on whether L2 learners also evidence comparable performance differences.

3.6 Measures to encourage reliability

In a recent L2 processing study, Juffs and Harrington (1995) attempted to support the results of a previous (though subsequently published) study conducted by White and Juffs (1998). However, when some of the findings of the two studies differed in unexpected ways, Juffs and Harrington argued that such discrepancies could be due to the differing numbers of tokens in syntactic...
studies types across the studies. Similarly, differences between the findings of the White and Juffs study and those of Schachtner (1990) and Johnson and Newport (1991), on which the former research was based, are accounted for by differences in methodology (participant selection, task modality, and item selection) (Juffs and Harrington 1995:50).

Such explanations don’t inspire much confidence in the interpretation of L2 research results. In light of problems like these, it is clearly important to replicate studies using matched research designs, as well as to use multiple tasks within a study to increase reliability, as discussed earlier. Multiple tasks and cross-task comparisons per participant help ensure that findings are not artefacts of particular procedures or stimuli. In addition, both production and perception tasks should be used within a study for the reasons discussed above.

We can also learn from the experience of psycholinguists who insist on an adequate number of participants in control and native speaker groups, tokens carefully balanced across stimulus types, and an experimentally-determined number of fillers included in every test to obscure the researcher’s focus. Crucially, we need to do more controlled longitudinal studies with experimentally-elicited data from single learners over time. However, until such studies are supported longitudinally by other learners, and therefore become generalisable, we are limited in the claims we can convincingly make from reported results.

The studies in this volume contribute to L2 research in important ways. Their methodologies are a reflection of the continuing development and refinement of our field.

4. Overview of the Book

The papers following this introductory chapter are divided into two sections, Current Perspectives and New Directions in Generative L2A Studies.

The first of these sections opens with Cairns’ comments (Chapter 2) on methodological issues underlying first and second language acquisition research. Drawing from her extensive work in FLA, Cairns illustrates the importance of methodological clarity in experimental research, thus further underscoring the points made in this chapter.

The next three chapters in this section investigate parameter setting with respect to the acquisition of functional categories: Philippe Prévost (Chapter 3) adapts Platzack and Holmberg’s (1989) finiteness operator parameter to Shlonsky’s (1996) split CP analysis for German and tests whether L1 Spanish learners of L2 German can reset parameters for verb movement to AgrO. Silvia Montrul (Chapter 4) tests whether the acquisition of the cative clitic in L2 Spanish triggers the functional projection AgrIOP, following Lightfoot’s (1991) analysis of language diachrony. In Chapter 5, Licona, Diaz and Maxwell investigate null subjects in L2 Spanish from a relatively new perspective within parameter resetting studies: They combine full UG access, L1 transfer, and syntactic restructuring of intergrammars to argue that parameter-setting does not actually take place in L2 acquisition. Target-like outputs are analyzed instead as the result of grammar representations which, while retaining the L1 parameter setting, apply certain operations made available by UG in order to accommodate L2 input which is not compatible with that setting. The section ends with Pérez-Leroux and Li’s chapter (6) which provides a reanalysis of a previously attested phenomenon in L2 acquisition, namely the differential treatment of two types of island constraints (Noun Complement Clauses vs. Relative Clauses). Using Mazzirri’s Locality Theory they propose that the gradual acquisition of lexical aspects of the Case system is responsible for the delay in L2 learners’ ability to recognize extractions out of Noun Complement Clauses as ungrammatical.

The title of the next section of this volume, New Directions in Generative L2A Studies, suggests the nature of the papers to follow. Each reflects an area in the field which has not been a major focus of inquiry: syntax-semantics correspondences, L2 processing, code-switching from a UG perspective, and applications of Chomsky’s (1995) Minimalist Program to second language acquisition.

The first paper in this section is that of White, Brown, Bruhn-Guravite, Chen, Hirakawa, and Montrul, (Chapter 7) who investigate L2 learners’ access to UG thematic principles. Specifically, the researchers explore the acquisition of non-native argument structure — two types of verb phrases — by adult learners of English.

In the following chapter (8), Klein presents an overview of L2 processing studies and offers directions for future research, typified by the next paper. In Fernández’s Chapter 9, the researcher investigates the effects of language-specific processing routines and how L1 transfer in parasag L2 input might affect ultimate attainment in the non-native language.

Bhatia and Ritchie (Chapter 10) next provide an overview of code-switching (CS) studies, including recent research within a generative framework. In their second contribution Ritchie and Bhatia (Chapter 11) examine the patterning of
dummy verbs in several code-switching studies and argue that these data show C3 to follow from the Minimalist Program (MP), rather than to reflect a performance phenomenon, as proposed in an alternative hypothesis. Explanation from an MP perspective becomes the major focus of the next two chapters: While Yusa (Chapter 12) offers a new analysis of adjacency and ECP effects in the L2 English of Japanese speakers under the MP, Robertson and Soface (Chapter 13) examine whether speakers of German transfer their L1 pattern of V2 raising into English L2 and also find their results best explained by the MP.

To complete the acquisition picture reflected in this volume, the final paper poses a general learnability problem and its solution. Janet Fodor outlines the nature of a language learner's input and why learnability theory has given up the idea of "automatic triggering." She proposes an alternative in which a learner's parsing routines play a central role in detecting triggers from the underlying structure to set parameters. While not specifically conceived for L2 acquisition, this model has a number of implications for explaining differences between native and non-native grammars in terms of the potential parsing strategies the learner selects. The volume thus ends with a possible solution to some of the critical questions we have raised in this opening chapter, namely integrating what we have called "process issues" into the study of how L2 grammars develop.

Acknowledgments

The authors' names are listed in alphabetical order. Earlier versions of this chapter were presented at the CUNY Psycholinguistics Summer Seminar (Fin 1996) and at the L1/L2 Acquisition Seminar at MIT (Spring 1997). We thank Chuck Caru, Suzanne Flynn, Barbara Lust and Virginia Valian for insightful questions and comments All errors are our own.

Notes

1. While this list is certainly not exhaustive, much of the research that has been conducted in the field to date — including the majority of the chapters in this volume — speaks directly to one of these issues.
2. For a brief summary of these positions, see White 1996.
3. We use the term "intergrammar" to contrast it to "interlanguage". The former denotes the mental representation of the L2 in the learner's mind. The latter denotes the interfaces resulting from the intergrammar.

4. We are not here concerned with defending or motivating these premises, but refer the reader to references in psycholinguistics, first language acquisition and processing literature.
5. This is an empirical question, though not one that can be decided easily. For one thing, we believe that to settle this question (as, say, it can be settled), we need to do more than to look at purported "end states" and find divergences from native standards, since such divergences might be surface phenomena, rather than deriving from underlying grammatical differences.
6. For similar arguments and discussion, not necessarily restricted to L2 acquisition, see Carroll 1996; Hale 1996; Epstein, Flynn and Martohardjono 1999b; Fraxier and de Villiers 1999; Marmot 1997.
7. The discussion here will center around empirical studies of syntax acquisition to reflect this volume's papers, but most questions could be extended to studies involving other realms of acquisition. Also, we do not attempt here to offer an exhaustive account of empirical research methodology; many excellent texts have done that before us, for example Segui and Shattuck Hulsey 1997 and Tanenhaus, Guss and Culbertson (1994).
8. We assume that researchers always attempt to meet the criterion of validity, i.e. ensuring that their hypotheses are falsifiable and that their tests logically test those hypotheses. The question of native speaker responses and how well, or poorly, they reflect analysis sets forth by linguists is an issue for both acquisition research. Researchers and theoreticians to ponder; much empirical data may, indeed, be a way the two realms can inform each other.
9. The lack of criteria for standardization of proficiency levels across institutions creates an obvious problem when attempting to interpret such terms as "intermediate," "high intermediate," etc. Some studies try to resolve this problem by administering normed tests such as the Michigan Test of English Proficiency, but until all researchers use the same evaluation measure(s), this will be an annoyance, if not a growing problem in L2 research.
10. For a brief but important discussion of the additional processing load placed on production, see Fodor (this volume).
11. It has been claimed that contextualized versus decontextualized tasks require different skills and often produce different outcomes (see Hecke 1995 and references therein). To respond to arguments that the former more closely taps learners' internal grammar, researchers such as Robert and Soface (1992) provide stimulus sentences within natural contexts.
12. We must also distinguish the language-specific facts which need to be learned, assumed to be outside the realm of IBL parameter setting.
13. We make the assumption that experiments employing grammaticality judgments will have both ungrammatical and grammatical stimuli for reasons well-known in psycholinguistic research.
14. The potential reason for variability, in this hypothetical case, could be semantic, i.e. the -a-noun of the verbal complement. Of course, when there is no coherent pattern, this suggests "free variation," understood here to indicate that the learner does not yet have a measurable hypothesis or rule for a targeted structure.
15. Eickman (1994) reports conducting a binding study with different findings from those of Finer and Rosemior. However, he accounts for these differences by comparing his analysis which
examined individual results to that of Breselow and Finer who reported only group results.


References


Elaine Klein and Gita Martohardjono


Madden, D. Preston and L. Selinker (eds.), Variation in Second Language
Platzer, C. and A. Holmberg 1989. "The Role of AGR and Finiteness in
Germanic VO Languages." Working Papers In Scandinavian Syntax 43:
51–74.
Chicago: University Of Chicago Press.
(eds.) Linguistic Perspectives on Second Language Acquisition. Cambridge:
Cambridge University Press.
Talent in Second Language Acquisition." In L. Obler and D. Fein (eds.),
The Exceptional Brain: Neuropsychology of Talent and Special Abilities. New
York: Guilford Press.
Schwartz, B. 1994. "Word Order and Nominative Case in
Nonnative Language Acquisition: A Longitudinal Study of L1 Turkish
German Interlanguage." In T. Hoekstra and B. Schwartz (eds.) Language
Acquisition Studies In Generative Grammar. Amsterdam: John Benjamins.
In L. Eubank (ed.), 1999. Point Counterpoint: Universal Grammar in the
Second Language. Amsterdam: John Benjamins.
Schwartz, B. 1995. 'Transfer' and L2 Acquisition of Syntax: Where Are We
Now?" Plenary Address, Second Language Research Forum, Oct. 1 1995, Cornell
University.
Schwartz, B. and L. Eubank 1996. "What is the 'L2 Initial State'?" Second
Language Research 12, 1–6.
Oxford University Press.
'Savant' Linguist." Linguistics 84, 315–51.
Sonae, A. 1993. "Incomplete vs. Divergent Representations of Unaccusativity In
Strozer, J. 1992. "Non-native Language Acquisition from Principled and
Parameters Perspective." In H. Campos and F. Marasco-Gil (eds.) Current
Linguistics 4, 142–165.
Linguistics 11, 392–400.
Tarone, E., S. Gass and A. Cohen 1994. Research methodology In Second-
Thomas, M. 1989. "The Interpretation of English Reflexive Pronouns by Non-
Thomas, M. 1989. "Universal Grammar and the Interpretation of Reflexives in
Vainikka, A. and M. Young-Scholten 1994. "Direct Access to X'-Theory:
Evidence from Korean and Turkish Adults Learning German." In T.
Hoekstra and B.D. Schwartz (eds.) Language Acquisition Studies In
Vallan, V. 1990. "Logical and Psychological Constraints on the Acquisition
of Syntax." In L. Frazier and J. De Villers (eds.), Language Processing and
Vallan, V., S. Aubry and J. Hoeflir 1986. "Young Children's Imitation of
Sentence Subjects: Evidence of Processing Limitations." Developmental
Psychology 32, 153–64.
Stranding by Second Language Learners and Parametric Variation." Second
Language Research 1, 15–46.
and E. Williams (eds.), Parameters and Linguistic Theory. Dordrecht:
Reidel.


Part II

Current Perspectives on Generative L2A Studies