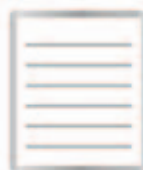


David Matley

EXPLORATORY GRAMMAR LEARNING

in a

MULTIMEDIA ENVIRONMENT



David Matley

**Exploratory grammar learning
in a multimedia environment**

David Matley

**Exploratory grammar learning
in a multimedia environment**

Tectum Verlag

David Matley

Exploratory grammar learning in a multimedia environment.

Zugl. Diss., Universität Tübingen 2011

Umschlagabbildungen: © stoyanh | shutterstock.de

Umschlaggestaltung: Heike Amthor | Tectum Verlag

© Tectum Verlag Marburg, 2012

ISBN 978-3-8288-5693-6

(Dieser Titel ist zugleich als gedrucktes Buch unter der
ISBN 978-3-8288-2950-3 im Tectum Verlag erschienen.)

Besuchen Sie uns im Internet

www.tectum-verlag.de

www.facebook.com/tectum.verlag

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte bibliografische Angaben sind
im Internet über <http://dnb.ddb.de> abrufbar.

To my parents
Philip Matley and Maureen Matley (née Gibbs)

Contents

Acknowledgements.....	11
Introduction	13
1 Changing approaches to second language acquisition research.....	15
1.1 The role of theory within second language acquisition research	15
1.2 The paradigm shift away from behaviourism.....	19
1.3 A permeable and dynamic system: interlanguage	26
1.4 'Learning' vs. 'acquisition' and the role of explicit knowledge	31
1.5 Universal Grammar approaches to L2 learning.....	36
1.6 Cognitive approaches to SLA	43
1.7 Conclusion	50
2 The role of explicit and implicit knowledge in L2 learning.....	55
2.1 Explicit vs. implicit knowledge	56
2.2 The interface issue: How do explicit and implicit knowledge interact?	67
2.3 Conclusion	95
3 Types of instruction and L2 grammar learning.....	99
3.1 Models of L2 instruction.....	100
3.2 Measuring the effectiveness of instruction	103
3.3 Results of empirical studies of explicit instruction	107
3.4 Conclusion	119
4 The role of the learner in L2 learning.....	123
4.1 Learning style.....	126
4.2 Learning strategies.....	133
4.3 Motivation.....	145
4.4 Conclusion	155

5	The role of the computer in L2 learning	157
5.1	Constructivist approaches to L2 learning	160
5.2	Hypermedia as cognitive tools for L2 learning	166
5.3	Learner autonomy and CALL	176
5.4	L2 learning tasks and CALL	185
5.5	Conclusion	195
6	The quantitative study	199
6.1	Pedagogical rationale and research questions	201
6.2	Participants and course design	203
6.3	Quantitative assessment	205
6.4	Data analysis	208
6.5	Discussion and conclusion	211
7	Description of the qualitative study	213
7.1	Aims of the study and research questions	214
7.2	Software and learning materials	216
7.3	Participants	225
7.4	Test methods	226
7.5	Analysis of cognitive processes	234
7.6	Concluding remarks	241
8	Use of subjects' existing explicit knowledge	243
8.1	Statement and elaboration of explicit knowledge	243
8.2	Dealing with uncertainty	259
8.3	Failure to notice a gap	276
8.4	Conclusion	300
9	Use of grammar explanations in the software	303
9.1	Confirming the subjects' own formulation	303
9.2	Noticing	315

9.3 Hypothesis formulation and testing	332
9.4 Conclusion	360
10 Use of other sources of explicit knowledge	363
10.1 Hypothesis formulation and testing	363
10.2 Use of explicit knowledge when using video excerpts	403
10.3 Conclusion	411
11 Summary and conclusion	413
11.1 The nature of learner knowledge of the L2.....	414
11.2 Learners' use of explicit knowledge	415
11.3 The impact of individual learner factors.....	419
11.4 The pedagogical relevance of multimedia CALL	420
11.5 Strengths and weaknesses of the study	424
11.6 Outlook for future research.....	426
Appendices	429
Appendix A: Assignment texts	429
Appendix B: Transcripts of video excerpts.....	431
Appendix C: Native speaker explanations of assignment forms	433
Appendix D: Learner profile data.....	436
Appendix E: Pretests and posttests in the qualitative study.....	438
Appendix F: Learner profile questionnaire (qualitative study).....	470
Appendix G: Abbreviations of tense forms	481
Bibliography	483
Index.....	525

Acknowledgements

In view of the time I have spent working on and finally completing this thesis, there are many people to whom I owe thanks. I would firstly like to thank all those people who accompanied me throughout the many years during which I worked on this PhD project and who experienced my various motivational highs and lows. They all know who they are, whether far away or close: thank you for bearing my frustration, bad moods and shouting at the computer, and for simply being there.

I am also indebted to several of my colleagues at the University of Zurich, who motivated me in many ways, one of which was completing their PhDs before I did. Additionally, I am grateful to the native speakers who gave their explanations of the grammatical forms in Appendix C. Finally, I would like to thank Prof. Dr. Kurt Kohn of the University of Tübingen for starting my interest in applied linguistics and for not giving up on me. I am grateful for his support and feedback as my thesis grew to fruition.

Zurich, February 2012

Introduction

This book deals with the field of computer-assisted language learning (CALL), with a particular focus on the use of explicit knowledge by learners working with multimedia grammar materials. The overall aim of my research project was to provide a better understanding of the ways in which learners make use of their explicit knowledge of the second language (L2) in a range of learning processes, in combination with an assessment of the impact of individual learner factors. It is based on the empirical research work that I conducted at the English Department of the University of Tübingen in Germany between April 1998 and July 2000.

Clearly, a great deal of time has elapsed from the beginning of my research work to the completion of this book. Conducting research and working as a Lektor at university was never easy; although being a practitioner of language teaching as well as a researcher has allowed me to gain many practical insights into what works and what does not work in the second language classroom, my experience also highlighted the difficulties that combining full-time teaching and research work involves. Nevertheless, despite the extended period of time over which this thesis has grown to fruition, it focuses on questions regarding the L2 learning process that are still of key relevance for teaching today: What cognitive processes do learners use when working with L2 learning materials? How do learners make use of the knowledge that they have of the L2? How can learners best be given support in autonomous learning? In what ways do individual learner differences influence the L2 learning process? How can computers best be used in the L2 classroom? It is these questions that this book addresses.

This book is divided into two main sections. The first section (Chapters 1 to 5) provides an overview of the theoretical and empirical insights from the fields of second language acquisition (SLA) and CALL that are relevant for the study. Chapter 1 conducts a resumé of the changing approaches to SLA research from its inception as a field of study. Chapter 2 focuses on the key issue of the nature of and interaction between explicit and implicit knowledge in L2 learning. Chapter 3 examines a range of models of the impact of instruction with a specific focus on L2 grammar learning. The focus of Chapter 4 is the role of the learner in the L2 learning process, covering the importance of factors such as motivation, learning style and learning strategies. The final chapter in this section (Chap-

ter 5) examines the role of the computer and multimedia as cognitive tools for L2 learning.

The second main section presents the results and discussion of the empirical studies carried out by the author. Chapter 6 describes the quantitative, comparative study of the effectiveness of teacher-based versus multimedia-supported grammar teaching. This is followed by the presentation of the findings of the qualitative, interpretative study of cognitive processes and use of explicit knowledge in Chapters 7 to 10, which forms the main research focus of this book. Chapter 7 describes the research design of the qualitative study, while the subsequent chapters analyse the use of the learners' existing explicit knowledge of the L2 (Chapter 8), of the grammatical explanations in the multimedia package (Chapter 9) and of other sources of explicit knowledge such as grammar exercises and video materials (Chapter 10). Chapter 11 then summarises the insights of the study, presents the key conclusions and makes recommendations for future research. A series of appendices provide a selection of teaching materials and measurement instruments such as pretests, posttests and learner profile questionnaires.

1 Changing approaches to second language acquisition research

This chapter gives an overview of the development of second language acquisition¹ (SLA) theory since its inception as a research discipline. Its aim is to place the current emphasis within SLA research on the language learner and on learning processes within a historical context. In particular, it assesses the potential of overall research perspectives within SLA for providing a theoretical framework for the focus of this study, namely the processes that L2 learners engage in when working with multimedia grammar materials.

Within SLA research there has been a growing realisation that the L2 learner is not a passive recipient of instruction; instead it has become widely accepted that the learner plays an active role in the learning process, to which they may bring idiosyncratic and indeed creative elements. This ‘emancipation of the learner’ calls for a research paradigm that can shed light on the learning process – one that can address L2 learning from a psychological rather than comparative linguistic perspective and explore what learners actually *do* when learning a second language rather than predicting what learners *should* do. It is these issues that the following chapter seeks to address, thus providing an overall framework for the subsequent chapters.

1.1 The role of theory within second language acquisition research

It is probably true to say that SLA research, some twenty-five years after its inception as an identifiable field of enquiry, is characterized by facts, opinions, explanations, positions, and perspectives that frequently exist in an uneasy state of complementarity and opposition. This is not to suggest that progress has not been made – in broadening the overall scope of

¹ Some researchers, notably Krashen (1977a, 1981, 1982, 1985), differentiate between ‘acquisition’ and ‘learning’, whereby the former refers to the subconscious learning of a language via exposure – ‘picking up’ a language – and the latter to conscious studying of the ‘rules’ of a language (cf. Krashen 1982: 10). In line with common practice, I will use these terms interchangeably to refer to the general process of learning a language in the broader sense, and will place them in single quotation marks if they are used in their technical sense. See R. Ellis (1994b: 14–15) and Section 1.4.

SLA research, in identifying the essential issues in need of investigation, in developing methods for studying them, and in collecting an enormous amount of data about them. Also, the discovery of competing and overlapping phenomena might be seen as evidence of the principal strength of SLA research – a willingness to explore a wide range of issues by means of alternative paradigms and methods. No doubt, over time, the pictures provided by the different sides of the prism will become clearer, but whether this will lead to a single, unifying account of L2 acquisition, as some believe to be necessary, remains to be seen. (R. Ellis 1994b: 689-690)

Second language (L2)² learning is a complex matter, a fact testified to by the sheer scope of Rod Ellis' 800-page *The Study of Second Language Acquisition* (R. Ellis 1994b) and the over 1000 pages of the second edition (R. Ellis 2008a). Since the 1950s and 1960s, research into second language acquisition has grown from a fledgling discipline into an established and extensive field of human study (cf. Gass & Selinker 1994: xiv; Mitchell & Myles 1998: 22). Research now covers areas ranging from naturalistic language acquisition to classroom-based learning, from the sociolinguistics of L2 learning to the role of universal grammar, and from the potential benefits of computer-assisted language learning (CALL) to L2 phonology.³ SLA is now a field with a broad spectrum of sub-disciplines, often with opposing assumptions and approaches, a field where the wide range of issues involved may not lend itself to one all-encompassing model. Ellis' scepticism is also echoed by Gregg (1984).

[S]econd language acquisition may simply be too difficult and too complex to be dealt with in a single theory. This seems like a reasonable idea, especially as Chomsky [...] has suggested the same sort of thing for first language acquisi-

² The research literature occasionally differentiates between learning a *second* language, where the language being learnt plays an institutional and/or social role in the community, such as French in Canada or Swedish in Finland, and learning a *foreign* language, where the language in question plays no such role and is primarily learned in the classroom, such as German in the UK or Japanese in Denmark. As R. Ellis (1994b: 12) points out, there is a need for an unambiguous superordinate term to refer to both forms of learning. In keeping with standard practice, however, I will use the term *second* language to cover both forms. Cf. R. Ellis (1994b: 11-12).

³ Interestingly, Rod Ellis' (1994b) monumental survey of second language acquisition research discusses all of these fields at length with the exception of computer-assisted language learning.

tion, where there are arguably fewer variables to worry about. (Gregg 1984: 79)

Yet despite the complexity and variety of second language learning processes, it can be argued that the field of SLA research has frequently been over-willing to adopt large-scale theoretical models and to uncritically take the validity of a theory for granted, Krashen's Monitor Model (Krashen 1981, 1982, 1985) being a case in point (cf. Gregg 1984; McLaughlin 1987: 55-58; Morrison & Low 1983).⁴ This tendency to quickly accept new 'alternative paradigms' and to conduct vast amounts of research in their name has in turn led to a "stifling of relevant research" (Schouten 1979: 4). Second language acquisition research has also often been considered to show "a tendency to build up from scratch, rather than building on what has already been established" (McLaughlin 1987: 12), and, in some critics' eyes, a weakness for faddism (McLaughlin 1987: 5; Nunan 1988: 174).

Not only has the field of SLA often all too eagerly implemented grand theory in its research work – SLA researchers and theorists have also tended to jump the gun in proclaiming that the 'insights' of their work should form the pedagogical basis for teaching languages. Yet language teachers have essentially been unimpressed by much of the potential pedagogical input offered by the field of SLA research, with audiolingualism being a case in point (R. Ellis 1990: 21). Teachers have tended to rely more on their own experience with learners in the classroom than on theoretical or empirical considerations from applied linguistics as a whole (cf. Borg 2009, 2010). Lightbown (1985) claims that "second-language acquisition research does not tell teachers *what* to teach, and what it says about *how* to teach they have already figured out" (1985: 182), whilst conceding that SLA research might help the teaching profession to have "more realistic expectations about what can be accomplished" (1985: 182). Other commentators go so far as to doubt

⁴ Krashen's Monitor Model (Krashen 1977a, 1981, 1982, 1985) is characteristic of both the benefits and drawbacks of an all-encompassing theory of SLA. On the one hand, its apparent simplicity and categorical nature provided the impetus for a vast amount of empirical research; on the other hand, much of Krashen's theory was seen to be seriously flawed by many researchers (cf. Gregg 1984; McLaughlin 1987: 55-58; Morrison & Low 1983), as far as both its practicability and internal consistency were concerned, thus putting into question much of the research conducted in its name. See also Section 1.4.

whether a *direct* interface between the teaching profession and the field of SLA research is possible at all (Widdowson 1990).⁵

However, since the mid-1980s, it can be said that the field of SLA research has begun to reflect more on its own development and shortcomings, and to be a little more wary of grand theory. Its attention has generally shifted away from all-encompassing macro-theories and towards micro-theories focusing on specific aspects of language learning, such as the role of instruction (R. Ellis 1990, 1994a) or the impact of psychological variables (Gardner 1985; Skehan 1989, 1991). From a more general perspective, there has been a gradual move away from a focus on *teaching* a second language – or in some cases, teaching learners to ‘unlearn’ their previous language ‘habits’ (cf. Skinner 1957) – towards an emphasis on *learning*, on language learners themselves, and the actual learning processes they undergo (cf. Kohn 1990: 1) This focus on learning processes has also been accompanied by increased interest in exploratory, qualitative methods (cf. Dörnyei 2007: 35-37; Richards 2003).

Thus the recurrent issues within the field of SLA have been those of methodology, the role and scope of theory, and the pedagogical implications of research outcomes, amounting to what can be described as ‘healthy debate’. The discipline of modern SLA research is still relatively young, and its status as an ‘art’ or ‘science’ is still in the balance (cf. Jordan 2004; Mitchell & Myles 1998: 193-194). For some, even the questions of the major accepted findings of almost fifty years of SLA research and of the key observable phenomena of the L2 learning process are still open ones (cf. R. Ellis 1994b: 679; Towell & Hawkins 1994).

However, although there may be disagreement as to what a ‘good universal theory’ of SLA has to account for, it is fair to say that there have been major insights into the complex and diverse process of language learning since the beginnings of the modern discipline in the post-war period (cf. Selinker 1992: 6). It is these key issues that will be reviewed in the remainder of this chapter. The following overview of key developments in the field of SLA will outline issues and concepts that are of relevance for this study by concentrating on the questions outlined below:

1. What is the role of the language learner in L2 learning?
2. What processes need to be understood in order to gain insights into L2 learning?

⁵ See also Christ (1998: 53-54) on the dangers of applying theories of one aspect of L2 learning to the pedagogy of teaching L2s as a whole.

3. What is the function of explicit knowledge, particularly in instructional L2 learning settings?
4. What role do learning strategies play in L2 acquisition?
5. What research paradigms are most suited to providing insights into the above questions?

As a result of this emphasis, attention will necessarily be placed more on instructed and adult L2 learning, rather than child SLA or L2 acquisition in naturalistic settings. Thus the overview makes no claims to be comprehensive: it will focus on core issues only. This brief outline of historical developments will provide the background for a more detailed discussion of the theoretical constructs that are central to this study in Chapters 2 to 5.

1.2 The paradigm shift away from behaviourism

Modern approaches to SLA research can be traced back to the work of Fries and Lado (Fries 1945; Lado 1957; cf. Kohn 1990: 2-3; Mitchell & Myles 1998: 23-24; Selinker 1992: 6-25), whose understanding of second language acquisition was anchored very much within the framework of a behaviourist theory of learning in general. Indeed, the behaviourist school of psychology dominated SLA research until well into the 1960s (Mitchell & Myles 1998: 25). Behaviourism essentially viewed language learning as the development of 'habits', an approach embodied in B.F. Skinner's *Verbal Behavior* of 1957. Skinner's work extended the notion of 'habit formation' from laboratory experiments on animals to the use of language by humans.⁶ Habits are formed when a stimulus is regularly linked with a response: a thirsty man, for instance, who is shown a glass of water which is just out of reach – the stimulus – will either reach for the glass or ask someone to pass it him by saying "Water" – the response

⁶ Skinner is essentially optimistic as to the validity of transferring results from animal behaviour to human verbal behaviour: "The basic processes and relations which give verbal behavior its special characteristics are now fairly well understood. Much of the experimental work responsible for this advance has been carried out on other species, but the results have proven to be surprisingly free of species restrictions. Recent work has shown that the methods can be extended to human behavior without serious modification." (1957: 3) He later adds: "We have no reason to assume [...] that verbal behavior differs in any fundamental respect from non-verbal behavior, or that any new principles must be invoked to account for it." (1957: 10)

(Skinner 1957: 32).⁷ According to behaviourism, a habit could be learned through both *imitation*, in which the learner copies the stimulus behaviour sufficiently frequently for it to become automatic, and – as in the above example – through *reinforcement*, whereby the learner is rewarded or punished depending on whether their response is appropriate or not – the reinforcement above being the quenching of the man’s thirst.

The behaviourists transferred this general model of learning to child first language (L1) acquisition, which they believed proceeded in exactly the same manner. Children would respond to stimuli in their environment, and imitate utterances produced by adults. Then, depending on whether the utterance was grammatically well formed or not, they would either be rewarded – reinforcement – or given correction (Skinner 1957: 29-30). Through repeated imitation and reinforcement, children would gradually develop a set of linguistic habits that would eventually constitute the structure of their first language.

The behaviourist school believed that second language acquisition would proceed in a very similar way to first language acquisition. The learner would learn a new language by systematically practising and imitating the structures of the L2 in a series of isolated stimulus-response tasks such as grammar drills (cf. R. Ellis 1986: 21), and would be rewarded by the teacher – say, with an encouraging remark (reinforcement) – or corrected depending on whether the utterance was accurate or not. Thus the key to L2 learning was also habit formation, a process in which *teaching* (grammar drilling) played a key role (cf. Kohn 1990: 7).

The main concern of the behaviourist approach to second language learning was learner errors. Behaviourist theory predicted that learners would carry over their linguistic habits from the first language to the L2, and that errors would occur where there were differences between the two languages. The fundamental mechanism within L2 learning was thus *transfer*: transfer would be positive where two languages shared a similar structure to express a certain meaning, thus resulting in no errors, and negative – a source of error – where the L2 had a different structure (cf. Kohn 1990: 2-3; Mitchell & Myles 1998: 23-25; Towell & Hawkins 1994: 18). Essentially, *difference* would cause *interference*: *difference* was equated with *difficulty*. In Lado’s words:⁸

⁷ Skinner plays down the exact notion of the stimulus and focuses more on the consequences of the response (cf. R. Ellis 1986: 20-21).

⁸ Lado’s *Linguistics across Cultures* (1957), despite no explicit references to Skinner, is clearly in the behaviourist tradition. In Section 3, “Grammatical structure, a system of habits”, he states: “In practical terms we understand that the

The student who comes into contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his [*sic*] native language will be simple for him [*sic*], and those elements that are different will be difficult. (Lado 1957: 2)

Thus, errors were “the result of non-learning, rather than wrong learning” (R. Ellis 1986: 22), and were to be avoided at all costs, as they were symptomatic of habits that had not yet been learned. They were a by-product of interference from the L1 and not worthy of study in their own right.

Behaviourism was thus a theory that provided an account of why learners made errors in learning an L2 – namely due to transfer from the L1. The pedagogical tool that was used in an attempt to predict exactly what kinds of errors learners would make was the method known as *Contrastive Analysis* (CA), an approach that can be traced back to the work of Fries (1945), in which he called for a systematic comparison of the differences between languages in the hope of predicting where difficulties would lie and hence errors would occur: “The most effective materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner.” (Fries 1945: 9). This challenge was later taken up by Lado (1957), whose work was to form the foundations of the modern discipline of SLA research (cf. Kohn 1990: 2; Selinker 1992: 6). The methods used were based on the *Contrastive Analysis Hypothesis*, initially formulated by Lado (1957), which claimed – in its strong form – that all L2 errors could be predicted by analysing the differences between the L1 and L2.⁹ This assumption provided the basis for an extensive research programme analysing the differences between languages as a potential

use of a grammatical structure by a speaker depends heavily on habit.” (1957: 58) and later “We simply do not realize the strength and the complexity of the habit system we have acquired through all the years of daily use of our native language.” (1957: 58)

⁹ The strong form of the hypothesis has since been shown to be inaccurate and has few supporters today (cf. Ellis 1986: 23-24; Larsen-Freeman & Long 1991: 55-56). Its weaker form, proposed by Wardhaugh (1970) claims only to be diagnostic, in that it can be used to identify which errors arise from interference and which do not. Yet the rationale of conducting an *a posteriori* contrastive analysis has also been criticised: “It makes little sense to conduct a lengthy comparison of two languages just to confirm that errors suspected of being interference errors are indeed so.” (R. Ellis 1986: 24). See also Section 1.3 of this thesis.

source of error, and dominated thinking within the fledgling discipline of SLA throughout the 1960s and into the 1970s (cf. R. Ellis 1994b: 306-309; Kohn 1990: 4; Larsen-Freeman & Long 1991: 52-56).

Hence, from a Contrastive Analysis perspective, there was a clear link between research and language pedagogy. Applied linguistic studies of differences between languages would be able to inform teachers where difficulties lay; languages were compared so that teachers would “know better what the real problems are and can provide for teaching them” (Lado 1957: 2). Learners, in turn, would be taught to avoid these difficulties by ‘breaking the habits’ their L1 had instilled in them. Teachers were thus the driving force behind language learning, whereas learners were passive recipients of drills and instruction. Language learning was essentially a question of ‘practice makes perfect’, with little scope for creative input from the learner.

This view of language learning, however, soon came under attack from both an empirical and a theoretical angle. The main weakness of CA was its lack of empirical support.¹⁰ Although Lado had called for research-based validation of the Contrastive Analysis hypothesis (Lado 1957: 72), there were few initial attempts to verify its accuracy (cf. Kohn 1990: 3). It was not until the early 1970s that a series of studies (Dulay & Burt 1973, 1974; George 1972; Grauberg 1971; for an overview see R. Ellis 1986: 28-30) began to show that many of the errors that learners were making in classrooms were not attributable to interference (negative transfer) from the L1. Specifically, native speakers of different languages were shown to be making similar types of error, ones that resembled the developmental errors made by child first language learners. Equally, not all areas of difference between languages actually led to difficulty. For example, when learning English word order, French learners were shown not to go through a stage of placing the object pronoun in front of the English verb, such as **I them like*, which CA predicted they should, as French places clitic pronouns before the verb as in *Je les aime* (cf. Mitchell & Myles 1998: 29-30; Towell & Hawkins 1994: 18-19; Zobl 1980). Thus the Contrastive Analysis Hypothesis was eventually shown to have failed in that it both over-predicted the errors that learners make – not all differences lead to errors – and under-predicted them – learners also make additional, developmental errors that do not arise from negative transfer from the L1. Crucially, what CA revealed was that a purely linguistic analysis of learner language would not lead to an explanation of the

¹⁰ See R. Ellis (1986: 27-33) for an assessment of theoretical, empirical and practical criticisms of CA.

source of learner errors or of the nature of learner language itself: a psychological explanation was required in its stead.

This lack of empirical support seriously undermined the basis of Contrastive Analysis. Yet more radical criticism was levelled at its theoretical origins in the behaviourist view of language learning. This critique was undertaken by the proponents of a mentalistic model of language learning, primarily in the work of Noam Chomsky and most notably in his 1959 review of Skinner's *Verbal Behavior*, in which Chomsky attacked the fundamental premises of Skinner's view of language learning. Skinner's assumptions about language learning were as follows:

- (a) The child language learner imitates what it hears.
- (b) The child's attempt to imitate is recognised by its parents and rewarded (*reinforced*).
- (c) The child wants to continue to be rewarded and thus repeats its attempts to imitate, resulting in the development of language-using habits.
- (d) All verbal behaviour (not only that of children, but also of adult L2 learners) is conditioned by such stimulus-response patterns.

Chomsky's counter-claims rejected all of these principles: if the basis for language use were imitated behaviour alone, children would not be able to produce the infinite number of sentences that they do, including ones they have never heard before. This was only possible, in Chomsky's view, if child L1 learners were born with the innate ability to generate abstract rules of the general structure of language (i.e. develop linguistic *competence*) from exposure to limited examples of actual language use. Furthermore, children go through similar developmental stages in acquiring linguistic rules, despite being exposed to different data. Chomsky also asserted that it is not the case that children only learn a language through 'meticulous care' on the part of adults or through reinforcement and correction of incorrect sentences. Rather than paying attention to correct grammar, parents focus primarily on meaning when engaging in communication with their children.¹¹

¹¹ Research has suggested that the primary focus of child-directed speech (CDS) is meaning rather than form; CDS is "semantically contingent" (Matychuk 2005: 316) in that it is meaningfully related to what the child says rather than on the form of the child's utterance. Nevertheless, some studies have suggested that parents provide subtle corrective feedback on their children's utterances in the form of modification or requests for clarification (cf. Mueller Gather-

Essentially, in Chomsky's view, the nature of language learning is fundamentally different to Skinner's model, and any accurate model must accommodate the actual phenomena of that process:

The child who learns a language has in some sense constructed the grammar for himself [*sic*] on the basis of his observation of sentences and nonsentences (i.e. corrections by the verbal community). Study of the actual observed ability of a speaker to distinguish sentences from nonsentences, detect ambiguities, etc., apparently forces us to the conclusion that this grammar is of extremely complex and abstract character, and that the young child has succeeded in carrying out what from the formal point of view, at least, seems to be a remarkable type of theory construction. Furthermore, this task is accomplished in an astonishingly short time, to a large extent independently of intelligence, and in a comparable way by all children. Any theory of learning must cope with these facts. (Chomsky 1959: 57)

Most importantly, however, Chomsky saw any attempt to speculate about the causation of 'verbal behaviour' as futile until the unique character of language was better understood, describing Skinner's speculations as "hopelessly premature" (1959: 55).

It is futile to inquire into the causation of verbal behavior until much more is known about the specific character of this behavior; and there is little point in speculating about the process of acquisition without much better understanding of what is acquired. (Chomsky 1959: 55)

Thus the nature of language itself must be understood before looking at what causes 'verbal behaviour'.

In an attempt to understand the rapidity and uniformity of first language acquisition, Chomsky posited the Language Acquisition Device (LAD), an innate pre-programmed system for acquiring language which generates abstract rules from specific instances of language use:

The fact that all normal children acquire essentially comparable grammars of great complexity with remarkable rapidity suggests that human beings are somehow specially designed to do this, with data-handling or 'hypothesis-formulating'

cole & Hoff 2007: 111). See Section 1.5 for a more detailed discussion of the role of input in Universal Grammar accounts of language learning.

ability of unknown character and complexity. (Chomsky 1959: 57)

Thus rather than being a procedure of pigeon-like habit formation, language learning is seen to be a distinctly creative process whereby some innate linguistic awareness is used to formulate and develop a complex language system.¹² This mentalistic model raised the question of the ‘true’ nature of language learning and led to a wealth of research on child language acquisition. From the 1960s on, Chomsky’s insights and the linguistic theory of Universal Grammar that lay behind it gradually became the predominant model of first language acquisition, leading to a shift away from and the eventual demise of behaviourist approaches to both L1 and L2 learning.

The impact of the paradigm shift on the study of L2 learning brought about by the demise of behaviourism was immense: the ultimate consequence of this change in thinking was a radical reassessment of the nature of the learning process, of the role of the learner and of the research methods needed to understand them. L2 learning was no longer seen as a simple matter of imitation and habit learning, but as a much more complex and creative process. Indeed, after the shortcomings of CA became apparent, researchers tentatively began to turn their attention to the whole notion of the *learning process* as such and the necessity of examining how learners processed input.¹³ CA had made the mistake of depending “solely upon an analysis of a *linguistic product* to yield a meaningful insight into a *psycholinguistic process*, i.e. second language learning” (Long & Sato 1984, cited in Larsen-Freeman & Long 1991: 56). The mentalistic approach and the shortcomings of CA had brought to light the misconceptions regarding what language learners actually do when they learn a language. Second language learners were clearly ‘up to something’ other than just transferring structures from their L1, and now attention began to turn to discovering what that ‘something’ was.¹⁴

The move away from behaviourism can also be seen as the first step towards an emancipation of the learner within SLA research paradigms. The focus shifted away from instruction and instead placed the learner at

¹² The nature of the LAD is discussed in more detail in Section 1.5.

¹³ However, even after the demise of CA, the initial focus within its ‘successor model’ Error Analysis (see Section 1.3) remained on the linguistic product, i.e. learner language and errors, rather than the learning process itself (Kohn 1990: 20).

¹⁴ However, the issue of *transfer* from the L1 has remained a central question within SLA research (cf. Odlin 2003).

the centre of research interest, a development that was to be fostered by the work of Corder (1967) and Selinker (1972). The learner had moved from playing a passive to an active and highly creative role in the learning process (Kohn 1990: 7), no longer merely functioning as the recipient of instruction. This development in turn had repercussions for teaching methodology: if the structures of a second language were not 'habits' that had to be drilled into the learner by the teacher (and old habits drilled out), then clearly the role of instruction had to be rethought. Now, at least from the viewpoint of a *theory* of language pedagogy, the issue of what effective instruction was remained wide open.

Importantly, too, theoretical considerations and *a priori* models of what learners produced when learning a second language had been shown to be out of touch with the actual linguistic structures that learners used, bolstering Lado's dictum that research claims should be verified against actual learner data: "[T]he list of problems resulting from the comparison of the foreign language with the native language [...] must be considered a list of hypothetical problems until final validation is achieved by checking against the actual speech of students." (Lado 1957: 72). This is not to say that grand theory was never to return to SLA research (cf. Section 1.4); however, the weaknesses of an all-encompassing model of SLA – and one imported from an external discipline at that – had become apparent. The focus of the work of researchers such as Corder (1967) and Selinker (1972) and of Error Analysis as a whole was thus to take an approach more clearly rooted in the research-then-theory tradition.

1.3 A permeable and dynamic system: interlanguage

It is not surprising that people holding the habit formation theory of learning, which has been the most prevalent theory over some decades now, showed no particular interest in the study of the learner's idiosyncratic sentences. They were evidence that the correct automatic habits of the target language had not yet been acquired. [...] The alternative view would suggest that the making of errors is an inevitable and indeed necessary part of the learning process. The 'correction' of error provides precisely the sort of negative evidence which is necessary to discover the correct concept or rule. (Corder 1971: 169-170)

During the 1960s and 1970s it had become clear that CA was unable to predict learners' errors accurately. Nevertheless, learners still made errors, but if they were not – at least not exclusively – determined by the

L1 or L2, then what lay behind them? In an attempt to answer this question, many researchers began examining learner errors in more detail, also focusing on non-contrastive ones. Although this method, labelled Error Analysis (EA), did have a tradition before Contrastive Analysis (e.g. Sridhar 1981; cf. Ellis 1986: 51), it was not until disillusionment with the usefulness of CA began to grow that it became a major research focus.

Overviews and critiques of the Error Analysis project as a whole have been given elsewhere (R. Ellis 1994b: 47-72; Long & Sato 1984; Richards 1974; Robinett & Schachter 1983); here suffice it to say that the main difficulty of EA was its scope: by focusing on errors and neglecting what learners do correctly, it failed to give a complete picture of learner language as such (cf. R. Ellis 1994b: 67-68). More importantly, it did not take account of *avoidance strategies* – the fact that learners may avoid using a certain structure of the L2 if they find it difficult (cf. Schachter 1974). The empirical insights it gave into *how* learners use language were limited. The truly “seminal”¹⁵ insights of this time were provided by Corder (1967) and Selinker (1972).

The work of Corder (1967, 1971, 1974, 1981) played a key role in the move towards a reassessment of learner language. In his view, learner errors had been much maligned and the dismissal of errors by CA had been a step in the wrong direction: “It almost seems as if they are dismissed as a matter of no particular importance, as possible annoying, distracting, but inevitable by-products of the process of learning a language about which the teacher should make as little fuss as possible.” (Corder 1967: 19)¹⁶ Corder also saw evidence from L1 learners that errors had a crucial role to play in the acquisition of a language:

[T]he best evidence that a child possesses construction rules is the occurrence of systematic errors, since, when the child

¹⁵ Selinker (1992: 150) uses the adjective of Corder (1967); R. Ellis uses it to describe both papers (1986: 50; 1994b: 48).

¹⁶ Corder was also critical of the potential pedagogical benefits of CA: “Teachers have not always been very impressed by this contribution from the linguist [Contrastive Analysis] for the reason that their practical experience has usually already shown them where these difficulties lie and they have not felt that the contribution of the linguist has provided them with any significantly new information. [...] [M]any of the errors with which they were familiar were not predicted by the linguist anyway. The teacher has been on the whole, therefore, more concerned with *how* to deal with these areas of difficulty than with the simple identification of them, and here has reasonably felt that the linguist has had little to say to him [*sic*].” (1967: 19-20)

speaks correctly, it is quite possible that he [*sic*] is only repeating something that he has heard. [...] It is by reducing the language to a simpler system than it is that the child reveals his [*sic*] tendency to induce rules. (Corder 1967: 22-23)

In the case of L2 learners, errors were equally (or arguably more) significant: learner errors could be used by the teacher to assess how far the learner had progressed and what they still needed to learn; they provided the researcher with an insight into the strategies and procedures that L2 learners used when discovering the language (Corder 1967: 25). Additionally, for learners themselves, errors were part of a crucial hypothesis-testing strategy:

Thirdly, (and in a sense this is their most important aspect) they are indispensable to the learner himself [*sic*] because we can regard the making of errors as a device the learner uses in order to learn. It is a way the learner has of testing his hypotheses about the nature of the language he [*sic*] is learning. (Corder 1967: 25)

This was a key reassessment of the nature of errors and of the learning process, one which recast what behaviourists considered to be 'interference' as the driving force behind L2 learning. In Selinker's words:

[Corder's contribution was] the insight that reframed our conception of 'errors' from something negative showing lazy unmotivated students to something normal and important for learning to occur, a 'window' on the learner's internal grammar, a learning strategy perhaps necessary to promote SLA. (Selinker 1992: 144)

Indeed, the notion of errors as strategies placed Corder's paper clearly within the mentalist tradition that had so scathingly criticised the entire behaviourist project.

Corder's contribution to an appreciation of the importance of learners' errors was expanded on and consolidated by the work of Larry Selinker. His 1972 article elaborates on the role of strategies and learning processes and the insights they provide. He lists five "central processes" (Selinker 1972: 216-217) in the process of learning the L2:

- (1) *language transfer* – structures are transferred from the L1 to the learner's linguistic system;
- (2) *transfer of training* – a rule is integrated by the learner as a result of instruction;

- (3) *strategies of second language learning*, which Selinker refers to as “an identifiable approach by the learner to the material to be learned” (1972: 216);
- (4) *strategies of second language communication* – “an identifiable approach by the learner to communication with native speakers” (1972: 217);
- (5) *overgeneralisation of linguistic material* – i.e. the learner overapplies the grammatical and semantic features of the target language.

These five processes together constitute the way in which the learner tries to internalise the L2 system (cf. R. Ellis 1986: 48).¹⁷

Selinker also discusses the question of whether – or which – learners have access to the innate language learning faculty of the mind posited by Chomsky. Selinker claims (1972: 212) that the “5%” of learners who are successful in learning a second language must go through different psycholinguistic processes than the vast (unlucky) majority, who fail to attain anything like mastery of the L2.¹⁸ The former, in Selinker’s view, still have access to the Language Acquisition Device (LAD) put forward by Chomsky – in Selinker’s terms the *latent language structure*¹⁹ – whereas the latter must have recourse to general cognitive mechanisms – *latent psychological structure* (1972: 212). Thus SLA could theoretically proceed in two different ways. In postulating this, Selinker had pointed out the key issue of the *route* of SLA and the problem of fossilisation – the fact that at some stage many if not most L2 learners cease to make any progress in learning the L2 regardless of how much input they are exposed

¹⁷ Although the ground-breaking work of Corder (1967) and Selinker (1972) picked up on the importance of the language learning process, their calls for a focus on processes remained largely ignored until the 1990s (cf. Kohn 1990: 20 *et passim*).

¹⁸ Selinker adds that these 5% of successful learners “may safely be ignored” (1972: 212). Selinker’s numerical estimate has been criticised as inaccurate, whereas the claim that successful L2 learners constitute a ‘pathological’ minority of no relevance for L2 research (cf. Bley-Vroman 1989, 1990) has been described as a ‘myth’ (cf. Herschensohn 2000: 46–48).

¹⁹ Selinker borrows the term *latent language structure* from Lenneberg (1967). According to Lenneberg the *latent language structure* “(a) is an already formulated arrangement in the brain, (b) is the biological counterpart to universal grammar, and (c) is transformed by the infant into the *realized structure* of a particular grammar in accordance with certain maturational stages.” (Lenneberg 1967, cited in Selinker 1972: 211–212, emphasis in original). Thus for all intents and purposes it can be seen as equivalent to Chomsky’s Language Acquisition Device.

to or how much they use their L2 skills in communication.²⁰ He had also raised a key question – in many researchers’ eyes *the* question in SLA – as to the role of implicit and explicit knowledge of the L2 by distinguishing between those few learners who have access to an innate system and the vast majority who have to integrate rules on the basis of instruction.

Both Selinker’s and Corder’s papers emphatically pointed out the central role of language learning strategies. Their initial description and elaboration of the role of strategies in the L2 learning process was a revelation of active learner-internal processing of language and the final break with the behaviourist notion of the role of errors and Contrastive Analysis (cf. R. Ellis 1986: 47-48). Making errors had now been reanalysed as a clear hypothesis-testing strategy on behalf of the learner, whereby the learner “revises the interim systems to accommodate new hypotheses about the target language system” (R. Ellis 1986: 50). Hypothesis testing had become an essential component of the development of a dynamic system. In addition, the five strategies identified by Selinker were an important first step towards a growing understanding of a wide range of learner strategies to be developed by later researchers (cf. O’Malley & Chamot 1990; Oxford 1990, 2011; Wenden & Rubin 1987).

Learner language was also shown to be systematic in nature; the structures L2 learners used were “idiosyncratic dialects” (Corder 1971), an ‘interlanguage’ (Selinker 1972) in its own right; learner language as a whole was seen to be not maverick, with ‘wild’ grammars, but rule-based. “[The learner] does not select haphazardly from his [*sic*] store of interlanguage rules, but in predictable ways. He bases his performance plans on his existing rule system in much the same way as the native speaker bases his plans on his internalized knowledge of the L1 system.” (R. Ellis 1986: 51) Both papers saw parallels between L1 and L2 learning

²⁰ The term *fossilisation* is generally used to refer to when the learner “ceases to elaborate the interlanguage in some respect, no matter how long there is new exposure, new data, or new teaching” (McLaughlin 1987: 61). This is regardless of how ‘closely’ the learner approaches native-speaker-like knowledge of the L2. The term *incomplete success* is used to emphasise the fact that L2 learners rarely achieve native speaker levels of L2 competence (cf. Bley-Vroman 1989: 43; Mitchell & Myles 1998: 13). However, in view of the worldwide growth of English as a lingua franca, the orientation within English teaching towards the native speaker as the sole benchmark of accuracy has been questioned by several researchers (cf. Jenkins 2005; Kohn 2007; Seidlhofer 2005). The implications of English’s role as a world language for teaching are discussed in Kohn (2007).

in that each could be separated into stages, each with its own logical internal structure (cf. Corder 1971).

These insights thus saw learner language as *permeable* (interlanguage rules are not fixed but open to continual amendment and revision), *dynamic* (interlanguage is undergoing constant change) and *systematic* (interlanguage grammars are not wild but rule-based). They suggested that SLA was a creative process in which the learner was actively involved through the use of learning and communication strategies,²¹ and in which both explicit and implicit knowledge could play a role. This in turn raised many questions regarding the nature of effective instruction: How do implicit and explicit knowledge interact in L2 learning? If the learner is actively and strategically involved in the learning process, then what kind of teaching approaches are most appropriate? Furthermore, if there is a dual route to SLA – via general cognitive mechanisms or via access to implicit language knowledge – how could instruction make use of the benefits of a ‘natural’ route to learning? One answer to the latter question was seen to lie in ‘communicative’ language teaching approaches, outlined in Section 1.4 below.

1.4 ‘Learning’ vs. ‘acquisition’ and the role of explicit knowledge

Error analysis involved a large number of investigations into the actual characteristics of learner language. In the 1970s a large number of so-called morpheme studies (cf. Bailey, Madden & Krashen 1974; Dulay & Burt 1973, 1974; Krashen *et al.* 1978; Larsen-Freeman 1976) were conducted in an attempt to examine the sequence in which L2 learners learned structures of the L2. Although the morpheme studies have been the subject of much criticism, in particular regarding their methodology (R. Ellis 1994b: 95-96; Gass & Selinker 1994: 84-87; Gregg 1984; Hatch 1978; Kohn 1990: 26), the general view that L2 learners go through developmental stages similar to those of L1 learners has now become widely accepted (Larsen-Freeman & Long 1991: 88-92; Mitchell & Myles 2004: 40-44; Towell & Hawkins 1994: 10-11).²² The work of Krashen in particular tried to

²¹ At the time of Selinker’s article, the concept of learning strategies was seen as “pure conjecture” (Richards 1974: 40). The nature of learning strategies is still a matter of debate within SLA research (cf. Grenfell & Harris 1998; Zimmermann 1997). See also Section 4.2.

²² Towell and Hawkins (1994: 5), for example, list “staged development” as one of the five key observable phenomena that any comprehensive theory of SLA must account for.

put many of the observations that Error Analysis and the morpheme studies made into the framework of an all-encompassing model of SLA. As Krashen's Monitor Model²³ has been described and criticised in detail elsewhere (Gass & Selinker 1994: 144-151; Gregg 1984; McLaughlin 1987: 19-58; Mitchell & Myles 1998: 35-39), this section will focus only on the key aspects of relevance for this study.

Krashen's Monitor Model, expounded upon a series of articles and books (Krashen 1977a, 1981, 1982, 1985) is based on five hypotheses, the most influential of which has been the Acquisition-Learning Hypothesis. Its key tenet is that 'acquisition' and 'learning' are distinct processes: 'acquisition' refers to the "subconscious process identical in all important ways to the process children utilize in acquiring their first language" (Krashen 1985: 1). Thus 'acquisition' is the product of exposure to the language (via communication) and involves developmental processes which involve some innate system similar (if not identical to) Chomsky's LAD. Krashen paraphrases 'acquisition' as "implicit learning, informal learning, and natural learning" and "picking-up [*sic*] a language" (1982: 10). 'Acquisition' is thus purely intuitive in nature. 'Learning', on the other hand, refers to a "conscious process that results in 'knowing about' language" (Krashen 1982: 10). It results primarily from explicit classroom teaching of the target language and from a focus on form. 'Learning' leads to explicit knowledge of rules of the L2 and is thus "formal knowledge, explicit learning" (Krashen 1982: 10).

Krashen also states that the innate language learning ability plays a key role in the acquisition process:

The acquisition-learning hypothesis claims, however, that adults also acquire, that the ability to "pick-up" [*sic*] languages does not disappear at puberty. This does not mean that adults will always be able to achieve native-like levels in a second language. It does mean that adults can access the

²³ The term 'Monitor Model' is used loosely – such as by McLaughlin (1987) and Mitchell & Myles (1994) – to describe the main hypotheses put forward by Krashen. It is alternatively referred to as 'Monitor Theory' (cf. VanPatten & Williams 2007a: 29-33). The *Monitor Hypothesis* – which posits that explicit knowledge can only be accessed during language production if learners are given sufficient time and only for the purpose of editing 'acquired' (i.e. implicit) knowledge (Krashen 1982: 15-20) – forms only one element of Krashen's theory of L2 learning. In line with other researchers, however, I will use the term 'Monitor Model' to refer to Krashen's framework of L2 learning as a whole.

same natural “language acquisition device” that children use.
(Krashen 1982: 10)

The key issue is the interaction between ‘learned’ and ‘acquired’ knowledge. Crucially, Krashen states that ‘learning’ cannot become ‘acquisition’: explicit knowledge of grammar rules gained through instruction in a classroom setting cannot make any contribution to the learner’s overall interlanguage system and cannot be incorporated into their language competence (cf. Krashen 1985: 83-124), in clear opposition to Selinker’s (1972: 37) view. ‘Learned’, explicit knowledge (that of grammar rules, for example) only plays a “trivial” role (Krashen 1985: 84) in that it can be used to monitor the learner’s output. In a similar vein, Krashen’s Natural Order Hypothesis (cf. Krashen 1982: 12-15) states that ‘acquisition’ takes place in a more or less fixed sequence, which is immune to the influence of teaching and ‘learned’ knowledge. Equally, Krashen’s Monitor Hypothesis (cf. Krashen 1982: 15-20) states that ‘learning’ can only be used as an ‘editor’, to make changes to the form of the learner’s utterances after these have been initiated by ‘acquired’ knowledge. Thus output is produced by ‘acquisition’ and then edited by the ‘learned’ system.

It is fair to say that Krashen’s model of language learning has been criticised more than any other theory of SLA, but that much of that criticism has been justified. Perhaps Krashen’s greatest failing was to present untested hypotheses as though they were a comprehensive model with empirical validity, and to use these hypotheses to draw pedagogical implications (Mitchell & Myles 1998: 39), which Krashen does very readily indeed, devoting extensive sections of his writings to the ‘pedagogical implications’ of his Monitor Model (cf. Krashen 1981: 100-118; 1982: 125-190; 1985: 69-92). From a theoretical point of view, Krashen’s model is also seriously flawed in that none of his hypotheses are falsifiable (Gass & Selinker 1994: 148-151; Gregg 1984: 94; McLaughlin 1987: 19-58; 1990a: 621), nor is there any effective empirical operationalisation of the concepts involved (Kohn 1990: 18; Schlak 1999: 8). In short, Krashen’s theory has been regarded as failing to meet the correspondence norms of a “good theory” (cf. Jordan 2004: 178-183; McLaughlin 1987: 12-14) and lacking essential theoretical precision (Kohn 1990: 18).²⁴

²⁴ The model is also accused of including redundant concepts and being self-contradictory (Gregg 1984: 84), atheoretical (Gregg 1984: 95) and so incoherent as to not warrant the term ‘theory’ (Gregg 1984: 94; Jordan 2004: 181). For a stringent criticism of Krashen’s work and its status as a ‘theory’ of L2 learning, see Gregg (1984) and Jordan (2004: 178-183). For a discussion of Krashen’s ap-

Similarly, by attempting to explain everything about the second language learning process, Krashen's all-encompassing model does not really explain anything in detail at all. It fails to pick up on the more explanatory approach of Corder and Selinker – regarding learning strategies, for example (Corder 1967, 1971; Selinker 1972) – and has little if anything to say about the processes involved in language learning as a whole (Kohn 1990: 18; McLaughlin 1987: 58). Nor does the model offer any insight into the precise nature of communication strategies (Selinker 1972) or the processing of input. In essence, the model has is regarded as too simplistic to represent the complex nature of linguistic performance (Kohn 1990: 19).²⁵

Despite these criticisms, Krashen's influence on the historical development of SLA research has been immense. Even though many view Krashen's claim that 'learning' cannot turn into 'acquisition' as being counter-intuitive, his emphatic distinction between the two types of knowledge has been very influential in producing and moulding research agendas, and in initiating a buzz of research activity which has led to insights in the field of SLA (Gass & Selinker 1994: 151; Mitchell & Myles 1998: 39). Krashen's theory emphatically placed the 'interface issue' of how implicit and explicit knowledge interact – a question touched on by Selinker (1972) – at the centre of research into SLA. Krashen's main contribution to the discipline has been to draw attention to the gap between 'learning' and 'acquisition' and to developmental sequences in SLA, and to offer a 'lay explanation' for why learner language is often so immune to instruction. Krashen rightly points out that there are only a relatively small number of rules that can be explicitly taught in comparison to the entire set of the grammar of English (Krashen 1982: 92-94), and that learners do manage to learn structures for which they cannot formulate rules. His call for more communicative teaching methods that moved away from a focus on form to a focus on meaning instead, providing 'comprehensible input' that would encourage the 'acquisition' of language rules, was also highly influential. Even

parent willingness to 'twist' evidence to support his claims, see Gregg (1984: 83).

²⁵ Gregg (1984) cites the example of a native speaker of Chinese with near-native English proficiency, yet who had not acquired 3rd person singular *-s*. Krashen's model would explain this incomplete knowledge as being due to the Affective Filter, but this offers no real explanation of why most of the L1 input would pass through the filter and only 3rd person singular would be filtered out.

McLaughlin (1987), who offers a highly critical appraisal of Krashen's theory, concedes that some of Krashen's ideas have their merits:

This is not to say that Krashen is wrong in his prescriptions about language teaching. Many researchers working in the field agree with him on basic assumptions, such as the need to move from grammar-based to communicatively oriented language instruction, the role of affective factors in language learning, and the importance of acquisitional sequences in second-language development. (McLaughlin 1987: 57)

Thus the lasting impact of Krashen's model of SLA has been to question the role of instruction and explicit knowledge in the language learning process and to focus debate on the implicit-explicit distinction. The learner of an L2 is seen to implicitly 'pick up' the language much in the same way as a native speaker learns their L1 – provided the L2 learner is exposed to the 'right' kind of input.²⁶ Thus in Krashen's view, the L2 learner is as creative and active as the L1 learner is. Yet his model crucially lacks any explanatory power whatsoever to convincingly suggest why this might be so. Krashen has nothing of value to say about language learning strategies, and offers no real understanding of L2 learning. His model dangerously trivialises the complex realities of both instructed and naturalistic language learning (cf. Kohn 1990: 19). In short, it leaves the 'black box' of language learning processes unopened.

Subsequent models of SLA have attempted to explore what the contents of that black box might be. In the majority of research work from the early 1980s onwards, the focus of debate thus turned to internal processing mechanisms. The central issue has been whether L2 learning is based on similar principles to L1 learning or whether an L2 is learnt in line with the acquisition of more general cognitive skills (cf. Mitchell & Myles 1998: 40). Indeed, Krashen's distinction between explicit and implicit learning ('learning' and 'acquisition' respectively) has lain at the heart of the matter. Recent approaches to this question can be divided into two broad theoretical perspectives. One is based on the Chomskyan linguistic tradition and advocates some role in SLA for an innate language learning faculty – Universal Grammar (UG) – and hence primari-

²⁶ However, Krashen's Input Hypothesis, namely that learners 'acquire' language by exposure to 'comprehensible input' that is just beyond their current developmental level (Krashen 1985: 2), has been subjected to fierce criticism, primarily on theoretical grounds: "The Input Hypothesis is untestable because no definition is given of the key concept, 'comprehensible input'." (McLaughlin 1987: 56)

ly, though not exclusively, implicit knowledge in the L2 learning process. The other, rooted in cognitive psychology, places a greater emphasis on explicit knowledge and cognitive processing. Cognitive approaches to SLA will be dealt with in Section 1.6. The insights of UG-based theory will be discussed in the following section.

1.5 Universal Grammar approaches to L2 learning

The focus on the naturalistic ‘acquisition’ of languages during the 1970s had led to the insight that there were similarities between the processes of L1 and L2 learning. The ‘discovery’ of developmental errors and cross-linguistic similarities in the sequence of acquisition of grammatical structures suggested that there was some evidence for a natural route to learning (Dulay & Burt 1974; Larsen-Freeman 1976; Krashen *et al.* 1978). Yet it was not until the publication of Chomsky’s *Lectures on Government and Binding* in the early 1980s (Chomsky 1981) that SLA researchers began to ask whether mentalist perspectives on L1 learning and the Universal Grammar model could be applied to the study of L2 acquisition (cf. R. Ellis 1994b: 440). From the 1980s onwards UG approaches became a major focus within SLA research, and have continued to dominate much thinking within the field (cf. White 2003b).

The key tenet of Universal Grammar is that all humans are born with a language learning mechanism, the Language Acquisition Device (LAD), consisting of an innate set of principles and parameters which determine the forms that languages can take. *Principles* are unvarying abstract properties of grammar that form the basis for the structure of all naturally occurring languages: all natural languages, for instance, are *structurally dependent*, whereby elements are not arranged in linear fashion but in hierarchical relationships to each other (cf. Mitchell & Myles 1998: 51-53). Whereas principles are uniform throughout all languages, *parameters* can have a number of different values which characterise individual languages. One instance of a parameter is *pro-drop* (Chomsky 1988: 64) or *null subject* (cf. White 2003b: 102-108), which determines whether a language can allow the deletion of subject pronouns: Spanish, for example is *pro-drop*, since it allows null subjects as in 1. below:²⁷

1. *Salieron a los ocho.* (*pro* left at eight)

²⁷ The examples are taken from Licerias (1989), cited in Towell and Hawkins (1994: 115).

Standard English, on the other hand, is a non-pro-drop language, since subject pronouns are required:²⁸

2. *They left at eight.*

The pro-drop parameter has two settings; others have multiple ones (cf. R. Ellis 1994b: 431-432).²⁹ These innate principles and parameters are seen to play an essential role in L1 acquisition.

Proponents of UG contend that children could not learn a first language as effortlessly and quickly as they do without the help of an innate, language-specific faculty equipped with these principles and parameters. On the basis of “degenerate input”³⁰ – the spoken language full of slips of the tongue, false starts, etc. (cf. Chomsky: 1965) – a child manages to make a mental representation of a first language that is far more complex than could be expected on the basis of mere input alone. “Children achieve this at an age when they have difficulty grasping abstract concepts, yet language is probably the most abstract piece of knowledge they will ever possess.” (Mitchell & Myles 1998: 44) Thus access to Universal Grammar is seen as a prerequisite for the task of L1 learning.

In one key respect, learners of a second language are faced with a similar problem to L1 learners: how to construct an abstract system from the rather fragmentary input to which they are exposed. Yet there are clearly large differences between L1 and L2 learning: whereas L1 learning is a rapid and effortless process leading to uniform *success* (Towell & Hawkins 1994: 58), L2 learning is – for the vast majority of learners – a slow and painstaking procedure which ultimately leads to varying degrees of

²⁸ Doubts have been raised as to whether English is in fact non-pro-drop; certain dialects and registers of English do allow subject pronouns to be dropped and native speakers have also been shown to omit initial subject pronouns in their speech to language learners (cf. Gass & Lakshmanan 1991).

²⁹ For extensive discussions of principles and parameters in relation to SLA see: Cook (1994), Towell & Hawkins (1994: 74-152) and White (2003b).

³⁰ The question of nature, purpose and function of ‘motherese’ or child-directed speech (CDS) in first language acquisition is a hotly debated one (for a review, see Matychuk 2005). Empirical studies have shown that CDS is not as ‘degenerate’ as Chomsky initially suggested, with only around 10% of utterances non-fluent or unintelligible (Cross 1977; Newport *et al.* 1977; cf. Mueller Gath-ercole & Hoff 2007: 110). Equally, parents have been shown to make a range of adjustments to their speech when addressing infants, making it syntactically simpler, and more limited in lexical complexity (cf. Snow 1995: 180). Yet it has also been pointed out that, even if caretaker talk is not ‘degenerate’, this does not exclude the possibility of an innate UG system (Snow 1986; cf. R. Ellis 1994b: 248-251).

failure, in that L2 learner competence stops short of native speaker competence in the vast majority of cases, and even in the case of exceptional language learners (cf. Sorace 2003).³¹ Additional crucial differences are that L2 learners already have a command of an L1 and are cognitively mature, in stark contrast to children learning their native tongue.

However, the most hotly debated issue has remained whether the above differences are *qualitative* or *quantitative* in nature, i.e. whether, despite the obvious contrasts between L1 and L2 learning, the procedures involved are fundamentally different or not (cf. Bley-Vroman 1989, 1990, 2009). Within UG-based research into L2 acquisition, the central question has thus been to what degree second language learners have access to the mechanisms of UG – if they do so at all. The literature distinguishes between a number of possibilities (cf. R. Ellis 1994b: 453-455; Mitchell & Myles 1998: 61-62; Skehan 1998: 78).

- a) *Full access hypothesis*: L2 learners still have full access to UG just as children do, but their different communicative needs lead them not to achieve complete mastery of the L2 (cf. Flynn 1987, 1996; Schwartz & Sprouse 1996).³²
- b) *Indirect access hypothesis*: L2 learners do have access to UG, but via their first language and thus through the parameter settings of that knowledge (cf. Clahsen & Muysken 1989; Schachter 1988, 1996).
- c) *Partial/modular access hypothesis*: L2 learners only have access to certain parts of UG; other parts are no longer available (cf. Hawkins & Chan 1997; White 1992).
- d) *No access hypothesis*: L2 learners do not have access to UG after a certain critical age, and thus have to resort to other, more general problem-solving skills (cf. Bley-Vroman 1989, 1990; Meisel 1991, 1997).

The vast majority of UG-based studies has attempted to explore the above hypotheses by examining whether interlanguages conform to the universal principles of UG, or whether L2 learners are able to notice that

³¹ The question whether near-native L2 learners' overall state of competence is *incomplete*, or *complete* and *divergent* from the L2 remains contentious (cf. Sorace 2003: 135).

³² A further possibility is the *dual access hypothesis* (cf. R. Ellis 1994b: 454-455), in which adult L2 learners still have access to UG, but also make use of general problem-solving skills. These skills 'interfere' with the UG module and in themselves are generally inadequate for mastering a language. For a discussion see Felix (1985). For the purposes of this overview, such approaches are subsumed under the full access hypothesis.

the parameters of the L2 are different to those of their L1 and hence reset them (cf. Herschensohn 2000; White 2003a, 2003b).

One of the strengths of a UG-based approach is that it provides a comprehensive and detailed theoretical background that allows researchers to formulate very precise hypotheses on the nature of L2 learning, with the consequences of 'parameter resetting' being a case in point (cf. Mitchell & Myles 1998: 70). Yet in terms of the empirical findings of UG-based research, the key question as to the extent to which UG still plays a role in L2 learning remains unanswered. Despite the appealing precision of hypotheses based on a UG model and the large amount of research into the effects of both naturalistic exposure and focused instruction, there have been few clear-cut results that would support any of the models of access to UG in SLA discussed above (cf. Dekydtspotter & Sprouse 2001: 1-2; Schachter 1996). Hopes that studies of parameters would provide evidence of hierarchical effects have not been fulfilled.³³ "Instead, while some predictions have been borne out, others have not, with discrepant results usually generating the need to provide post-hoc explanations with ever more complicated versions of the underlying linguistics theory." (Skehan 1998: 79) Even staunch defenders of the UG approach have conceded that "results are not totally unproblematic as far as claims for parameters in interlanguage grammars are concerned" (White 2003b: 148). Thus much of the empirical research has been inconclusive (cf. Mitchell & Myles 1998: 68).

The ambiguous nature of empirical findings has been coupled with theoretical issues. Clearly, UG is a strong theory, providing as it does a rationale for the general absence of 'wild' grammars in interlanguage development³⁴ – via continued access to UG principles – and an account of the phenomenon of transfer from the L1 and its limitations – due to parameter resetting (cf. Towell & Hawkins 1994: 7-10). Thus, in terms of its

³³ A number of researchers call for caution when claiming that UG is indeed at work in SLA on the basis of parameter-resetting research. "We have to be careful [...] not to draw hasty conclusions on the basis of evidence relating to one structure only, and we have to bear in mind that other explanations which do not involve UG might be possible, and have indeed been put forward." (Mitchell & Myles 1998: 64). Indeed, apparent parameter setting could also be influenced by issues such as perceptual saliency (Andersen 1983, 1984, 1990; Andersen & Shirai 1994; Slobin 1985) or implicit learning (Reber 1967; Reber, Allen & Reber 1999; Winter & Reber 1994).

³⁴ For a discussion of the absence or presence of wild interlanguage grammars, see White (2003b: 42-53).

explanatory and predictive power, and heuristic value, it fulfils many of the criteria of a “good theory” (Jordan 2004; McLaughlin 1987: 12-18).³⁵

However, UG has serious drawbacks as a model of SLA. The heuristic value of UG has arguably been compromised by the vitality of the theory itself, as the model is undergoing constant development and revision, with one approach being dropped in favour of another, with the rise of a minimalist approach to UG being a case in point (cf. Chomsky 1995; Herschensohn 2000). Thus, although theory development may be beneficial for the proponents of UG as a whole, this leads to problems in using that theory within SLA research:

The problem is that while researchers working within UG may be content with this state of affairs, ‘consumers’ who are interested not in developments in linguistic theory for their own sake, but instead in the explanatory value the account can provide in related areas, find themselves stranded since the version of theory they are diligently testing proves to be abandoned by UG researchers themselves. (Skehan 1998: 79)

Indeed, as the linguistic constructs that form the basis of UG are being revised, this also limits the extent to which the theory is falsifiable as an explanation of L2 learning. Gass and Selinker (1994: 130) point out that researchers positing access to UG as an explanation of L2 learning can always argue that the linguistic constructs underlying the research were not the ‘correct’ ones if confronted with data that does not support their theses. Equally, in his assessment of UG as a model of SLA, Jordan (2004: 162) also contends that changing theoretical constructs and ‘goal-post shifting’ make it difficult to falsify the theory. Thus the interface between UG theory and UG-based SLA research remains problematic and ultimately may undermine much of the research conducted in its name.

The difficult nature of the interface issue is also reflected in UG’s primary focus on the competence-performance distinction.³⁶ The crucial point

³⁵ Towell and Hawkins (1994: 57) see an additional strength of UG in that it is not an *ad hoc* theory that has been devised to account specifically for SLA, but one that attempts to explain the phenomenon of language as a whole.

³⁶ Further criticism has focused on methodological concerns within UG-based L2 research, in particular regarding the validity and reliability of using grammaticality judgements to assess L2 learner competence. Cook describes grammaticality judgements as “a source of evidence that has to be treated with extreme caution as it is unclear how directly it taps the individual’s knowledge of lan-

is that UG is primarily a theory of L1 competence and not necessarily one of L2 performance:

The more general issue [...] is how to make the connection from a linguistic theory of language competence to a theory of second-language learning. Chomsky is not concerned in his writings with second-language learning. The burden rests on those who would apply his ideas to second language to show how the connection is to be made. (McLaughlin 1987: 108)

UG is first and foremost about L1 learning, and hence should not be judged “in terms of the aims and objectives of a *second language learning theory*” (Mitchell & Myles 1998: 69).³⁷ Jordan (2004: 154) also concurs that UG is “too strict and too narrow” to be viable as a theory of SLA.

However, not only does it fall short in terms of providing an understanding of performance, the UG approach also provides little insight into language learning as a *process*. There has been a clear discrepancy between the focus of UG-based SLA research on the one hand – namely on learner language as a *product* and elusive L2 competence – and the approach Selinker called for in his 1972 paper – namely that SLA research should investigate the *processes* of L2 learning (cf. Kohn 1990: 20). Central constructs such as noticing and hypothesis testing are neglected within a UG model of L2 learning (cf. McLaughlin 1987: 107-108; Skehan 1998: 80). UG thus neglects key elements of the process of adult language learning:

[F]or older learners, meaning is primary, strategic language use is pervasive, and processing-based generalisations have considerable utility in accounting for second language performance. [...] UG is good at describing a formal, underlying competence (and possibly learning in the pre-critical period stage), but it is less convincing with second language learning, with real-time communication, and with the relationship between performance and change. (Skehan 1998: 79-80)

Thus, despite its theoretical strengths, the UG model “does not address how learners acquire the ‘skill’ of using their grammatical knowledge” (R. Ellis 1994b: 458). In taking a somewhat clinical view of language as

guage” (1994: 442). Cf. also R. Ellis (1994b: 441-442) and Sorace (1996). See also Section 3.2.

³⁷ Cf. White (2003a: 36): “Although UG provides constraints on possible grammars in the course of acquisition, it is not, of itself, a theory of acquisition.”

competence only, it neglects the vital social role that language plays (Mitchell & Myles 1998: 69), and indeed the active role that the learner plays in the L2 learning process. Jordan's (2004: 164) conclusion – based on the narrow scope of UG – is highly critical of its potential as an approach to understanding the key processes of SLA: "Even assuming that UG exists, that UG theories of L1 acquisition are true, and that L1 learners have at least some access to UG, most of the questions that concern SLA researchers remain unanswered, indeed they are not addressed."

Overall then, a UG-based theory of SLA does not make any claims to be an all-encompassing macro-theory accounting for all aspects of L2 development; its primary concern is the "nature of unconscious interlanguage knowledge" (White 2003a: 37), and thus not the processes of L2 learning but the state of the mental 'product' in the learner's mind. Nevertheless, it is this very focus on linguistic competence that leads it to neglect the strategic manner in which learners make use of their knowledge of the L2 (cf. Kohn 1990: 72-80).³⁸ Equally, factors such as the role of explicit knowledge, learning strategies, noticing, attention and hypothesis testing are also beyond the scope of UG,³⁹ key phenomena that need to be understood to provide a more general understanding of the L2 learning process as a whole. It is perhaps for this reason that UG models have increasingly been incorporated into modular theories of SLA, which assign significant roles to both cognitive processing skills and a UG module (cf. R. Ellis 1990, 1994a; Gass & Selinker 1994: 295-309; Towell & Hawkins 1994).⁴⁰ The actual *processes* of L2 learning are the

³⁸ On the basis of sentences that pose comprehension difficulties for native speakers, Kohn (1990: 72-80) highlights the way in which learner performance does not consist of the mere application of linguistic competence, but is characterised by processes and strategies that control the use of that knowledge. Learners also make use of both strategic and world knowledge in attempting to understand utterances. "Speakers use their knowledge in strategic ways, as comprehensively and precisely as seems necessary and sufficient to ensure communication." (1990: 74; my trans.)

³⁹ See Herschensohn (2000: 183-214), however, for a discussion of UG and learning strategies.

⁴⁰ UG theory also pays little attention to other broader aspects of language. Despite recent studies of L2 phonology, morphology and the lexicon (cf. White 2003b: 178-240), its main focus has been on syntax. It even falls short of providing an explanation of the acquisition of 'grammar' as such since "it concerns itself only with those aspects of grammar that fall within the 'core'" (R. Ellis 1994b: 458). Crucial elements of language use such as lexical development, pragmatic competence and discourse are not accounted for (1994b: 458).

focus of cognitive approaches, which will be discussed in the following section.

1.6 Cognitive approaches to SLA

While UG approaches assert that first language acquisition can shed at least some light on L2 learning, cognitive models look to more general aspects of how humans process information in an attempt to understand SLA. In contrast to UG theory, which focuses on the linguistic competence of an ideal native speaker, and is not primarily concerned with how learners access their linguistic knowledge in real time, cognitive research concentrates on the strategies learners use in producing or understanding language and the performance of language learners.⁴¹ The rationale for these approaches is based on the conviction that L1 and L2 learning are different phenomena. This view has been articulated in Bley-Vroman's Fundamental Difference Hypothesis (Bley-Vroman 1989, 1990, 2009), which considers the "fundamental [...] relatively apparent, large-scale characteristics" (Bley-Vroman 1989: 43) of adult L2 learning that distinguish it from child L1 acquisition. These are summarised below (cf. Bley-Vroman 1990).

1. *Lack of success*: The vast majority of foreign language learners do not achieve anything like native levels in the L2, whereas children inevitably achieve complete mastery.
2. *General failure*: Failure is the norm in L2 learning; the few instances of apparent success should be seen as 'pathological' cases that are not central to the field of second language acquisition theory.
3. *Variation in success, course and strategy*: Not only do adults generally fail to learn a foreign language, they also fail to varying degrees. Variation is central to adult second language learning, while children follow a much narrower path of development. "Clearly, a formal model of adult foreign language learning must allow many different 'grammars' to be arrived at." (Bley-Vroman 1989: 45)

⁴¹ However, performance is not the sole concern of cognitive approaches. "It would be wrong to suggest, however, that cognitivists are only interested in performance and not in competence. In fact, the line that they draw between the two (if they draw a line at all) is much more fuzzy than for theoretical linguists, and they see them as closely related and interacting aspects of the learning process, each driving the other forward in a dynamic tension." (Mitchell & Myles 1998: 73)

4. *Variation in goals*: Adults do not all have the same objectives in learning an L2 – some want to be grammatically perfect, others ‘merely’ to make themselves understood. Child L1 learners, on the other hand, “do not have the luxury of setting their own individual goals” (Bley-Vroman 1989: 46). Indeed the very notion of setting objectives is unknown in first language development.
5. *Correlation of age and proficiency*: The major factor that appears to determine success in L2 learning is age of exposure to the L2 (cf. Birdsong 2006; Johnson & Newport 1989; Newport 1990; Singleton 2005). There seems to be a critical age for native-like proficiency for L2 learners.
6. *Fossilisation*: Adult L2 learners often stabilise their linguistic system at a certain stage, after which development stops and is highly resistant to change, even by means of focused instruction. In child language acquisition the system remains flexible until complete mastery is achieved.
7. *Indeterminate intuitions*: Even highly advanced L2 learners appear to be unable to judge the grammaticality of L2 sentences unequivocally, whereas native speakers generally are.
8. *Importance of instruction and practice*: Adult learners seem to require instruction, whereas it has no obvious role in the linguistic development of children.
9. *Negative evidence*: Child language acquisition does not appear to make use of negative evidence; adult language learning appears to need it or at least find it useful.
10. *Role of affective factors*: L1 learning appears to be wholly independent of affective factors such as personality, motivation, attitudes or anxiety. In L2 learning, the impact of affect, although still unclear, has been shown to be significant.

Thus the mechanisms behind first and second language learning are seen to be qualitatively distinct as shown in Table 1.1 below (Bley-Vroman 1989: 51).

Child language development	Adult foreign language learning
A. Universal Grammar B. Domain-specific learning processes	A. First language knowledge B. General problem-solving procedures

Table 1.1: *Processes in child language development and adult foreign language learning*

Bley-Vroman's claims are not uncontroversial within SLA research (cf. Herschensohn 2000: 46-53). In particular, the roles of instruction and negative evidence in SLA are unresolved. Yet cognitive approaches to SLA have essentially been based on an assumption of the above points, and have thus undertaken to explain why, for example, there is variation in learning strategies, why fossilisation occurs, what the role of instruction, negative feedback and practice is, and what the role of affective factors may be, whereas UG-based approaches have focused on domain-specific factors such as parameter resetting.

Thus cognitive approaches view L2 learning as a process that is based primarily on general learning and problem-solving skills rather than an inherent and separate language learning faculty. Cognitive accounts of SLA share the view that L2 learning involves the use of the same cognitive systems – memory, problem solving, information processing, etc. – as the acquisition of any other complex cognitive skill. The mechanisms involved are not specific to language learning but are based on general cognitive accounts of learning. Within such an approach, the language learning process is seen to be restricted by human information processing abilities (R. Ellis 1994b: 390; McLaughlin 1987: 148).

Several models of L2 learning based on cognitive psychology have been put forward since the 1980s, with a key focus on interlanguage development, access to linguistic knowledge, and cognitive strategies. Models include the perceptual saliency approach (Andersen 1983, 1984, 1990; Andersen & Shirai 1994; Slobin 1985), the Multidimensional Model and Teachability Theory (Clahsen, Meisel & Pienemann 1983; Pienemann 1981, 1989; Wolfe Quintero 1992), and connectionist or parallel distributed processing (PDP) approaches (Broeder & Plunkett 1994; N. Ellis 2003; N. Ellis & Schmidt 1997; Gasser 1990; MacWhinney 1989; Sokolik 1990; Sokolik & Smith 1992). However, despite the insights that these models have brought about, they will not be considered in more detail here. Instead this section will focus on information-processing approaches to language learning (Anderson 1983, 1985, McLaughlin 1987, 1990b), as it is the constructs of those models that have formed the foundations of

many recent integrated and modular models of L2 learning (cf. R. Ellis 1990, 1994a; Gass & Selinker 1994: 295-309; Hulstijn 2002; O'Malley & Chamot 1990; Towell & Hawkins 1994; Skehan 1998).

The model put forward by McLaughlin (1987, 1990b) is based on Shiffrin and Schneider's theory of memory and cognitive processing (Schneider & Shiffrin 1977; Shiffrin & Schneider 1977). This theory sees memory as a collection of nodes that become 'complexly associated' through learning (McLaughlin 1987: 134). Linguistic knowledge is stored in separate modules: long-term memory (LTM) with a large capacity for information storage, and short-term memory (STM), which is of limited capacity. New linguistic knowledge is acquired by the learner selecting features of the linguistic input to pay attention to and then storing these features in short-term memory. Later the information is transferred into long-term memory.

A key distinction is made in this, and indeed other information-processing models, between *automatic* and *controlled* processing.⁴² Controlled processing is a temporary activation of memory nodes (McLaughlin 1990b: 115). It requires large STM capacity, and thus the activation of such processes is slower. "Controlled processes are easy to set up, alter and apply to novel situations but they are also inefficient because they require time for activation and use up available processing capacity." (R. Ellis 1994a: 85). *Automatic processing*, on the other hand, involves the "activation of certain nodes in memory every time the appropriate inputs are present" (McLaughlin 1987: 134). An automatic process uses a relatively permanent set of associations in long-term memory, and thus requires extensive training and practice to develop. Once an automatic process has been learned, it is difficult to alter or repress. It can be performed without placing demands on short-term memory.

Once a skill has become automatised, attentional control can be allocated to the next stage of learning, freeing controlled processing capacity which can focus on new information and more complex knowledge structures. It is necessary for simpler structures to become automatised before more complex ones can be integrated (Mitchell & Myles 1998: 86).

⁴² The automatic-controlled distinction is not uncontroversial. The empirical basis for the distinction rests mainly on performance on relatively simple tasks combining visual processing and memory search (subjects are presented with a set of items and then asked to say whether any members of the set appear in subsequently presented items). Automatic processing may not be free of attention in more complex tasks. Also, it is a descriptive rather than explanatory approach (cf. McLaughlin & Heredia 1996).

Controlled processing then lays down the stepping stones for moving to the next stage of learning. Thus the development of language learning involves constructing a set of well-learned automatic procedures, which can be carried out in parallel (Schmidt 1990: 136), so that learners' limited controlled processing abilities can be freed for new learning and spread over the various components and sub-components of a task:

To learn a second language is to learn a skill, because various aspects of the task must be practised and integrated into fluent performance. This requires the automatization of component sub-skills. Learning is a cognitive process, because it is thought to involve internal representations that regulate and guide performance. [...] As performance improves, there is a constant restructuring as learners simplify, unify, and gain increasing control over their internal representations [...] These two notions – automatization and restructuring – are central to cognitive theory. (McLaughlin 1987: 133-134)

Information-processing models of SLA thus see language learning as the transfer of information from short-term to long-term memory and the gradual automatization of controlled processes via repeated practice (McLaughlin 1990b: 115).

Anderson's ACT* model of L2 learning also posits a key role for automatization via practice (Anderson 1983, 1985). Yet Anderson's model differs from McLaughlin's in that it postulates three different memory capacities: working memory, in line with STM, and two forms of LTM, one for *declarative* and one for *procedural* knowledge respectively.⁴³ Declarative knowledge is knowing *that*, such as knowing the rule that regular verbs take third person *-s* in the present tense in English; procedural knowledge is knowing *how*, such as knowing how to actually form third-person singular in performance (cf. Mitchell & Myles 1998: 87-88). Similarly to the automatic-controlled dimension, knowledge can shift from the declarative to the procedural form through practice. Anderson divides the learning process into three stages (cf. Mitchell & Myles 1998: 88).

- *The cognitive stage*: A description of the procedure is learned.
- *The associative stage*: A method for performing the skill is worked out.

⁴³ The declarative-procedural distinction should not be seen as a binary opposition, but as the end points of a continuum. The importance of this distinction in cognitive accounts of SLA and its interaction with the automatic-controlled and implicit-explicit dimensions are discussed in more detail in Section 2.1.

- *The autonomous stage*: The skill becomes more and more rapid and automatic.

In this model, learning takes place by declarative knowledge becoming proceduralised. When this occurs, it can be accessed automatically without recourse to the limited capacity of working memory. New declarative knowledge can be attended to and can eventually proceed through the associative and autonomous stage (Mitchell & Myles 1998: 88). Fluency in L2 learning is thus gained on the basis of proceduralised knowledge that becomes subject to automatic control.⁴⁴

Information-processing models of L2 learning have been seen to account for several aspects of SLA. The model of cognitive control provides a rationale for why learners perform differently from task to task, as individual tasks vary in the amount of attention they require or allow, and also in where attention needs to be focused (R. Ellis 1990: 176). The controlled-automatic dimension also offers an explanation of the step-by-step (incremental) nature of language learning, whereby simpler structures need to be automatised before more complex ones can be acquired (Mitchell & Myles 1998: 86). Furthermore, restructuring may also account for variability in interlanguage and the temporary reappearance of more basic errors, manifested in U-shaped learning (McLaughlin 1990b: 118). Expert learners may also be more flexible in their ability to restructure rules and thus avoid common errors (Nation & McLaughlin 1986). Information-processing perspectives also offer a rationale for fossilisation. The fact that learners stop short of native-like competence and are unable to remove non-native structures despite correction and extensive exposure to L1 input may be due to controlled processes becoming automatic – and thus outside the attentional control of the learner – before they are native-like, “giving rise to a stable but erroneous construction” (Mitchell & Myles 1998: 86).

Thus cognitive approaches evidently place the learner in a completely different context to that of the behaviourists of the 1950s and 1960s, and indeed of many UG-based models. Within cognitive models of L2 learning, the individual is seen as “*active, constructive, and planful* rather than a passive recipient of environmental stimulation” (McLaughlin 1990b: 113, emphasis in original). Learning is an “active and dynamic process in which individuals make use of a variety of information and strategic modes of processing” (O’Malley & Chamot 1990: 217). Its emphasis is on mental events, on “finding a scientific means for studying the mental

⁴⁴ Anderson’s model has been applied to various modules of language learning, in particular learning strategies (O’Malley & Chamot 1990; cf. also Section 4.2).

processes involved in the acquisition and application of knowledge" (McLaughlin 1990b: 113). Thus the learner plays an active role in the learning process.

An approach of this kind also has a key focus on the learning processes involved and emphasises the role of strategies within that process (O'Malley & Chamot 1990: 19-20). The specificity and dynamic, process-oriented nature of the theory makes it possible to provide a more detailed view of language learning than other theories and also provides a mechanism for describing how language learning can be improved (1990: 19). Construction of knowledge and strategies for language use and development are central concepts within a cognitive theory of SLA.⁴⁵ Cognitive accounts of second language learning thus include an analysis of the strategies people use for thinking, remembering, understanding and producing language (McLaughlin 1987: 133-134).

However, an information-processing perspective is not without its shortcomings. It lacks concretely formulated hypotheses about the L2 learning process, a shortcoming which McLaughlin himself also concedes: "Cognitive theory does not represent a highly articulated theoretical position. There have been relatively few attempts to spell out with any degree of precision what the predictions of such a theory would be for second language learning." (1987: 150). A further drawback of the theory – especially in comparison to UG approaches – is the lack of precision in its terms. The notion of improvement through 'practice', for example, leaves open the question of what 'practice' as such might involve: the opportunity to use language in communication, or practicing specific rules in drills and exercises (R. Ellis 1994b: 391). Equally, the vagueness of other concepts such as 'restructuring' makes it difficult to formulate falsifiable hypotheses as to which structures would be restructured when (R. Ellis 1994b: 391), and the extent to which concepts such as controlled versus automatic processing can be empirically tested is also seen as questionable (cf. Jordan 2004: 212). The theory also has difficulty in explaining some of the key observable phenomena of second language acquisition, such as staged development and apparently natural acquisition sequences in L2 learning (cf. Towell & Hawkins 1994: 10-11).

More recent research based on a broadly cognitive approach to L2 learning has attempted to address some of these criticisms. Researchers such as DeKeyser (2007b) and Leow (2007) have attempted to give more pre-

⁴⁵ The compatibility of a cognitive approach to L2 learning with a broadly constructivist approach to learning as a whole is discussed in Section 5.1.

cise definitions of key concepts such as practice (cf. Leow 2007: 21-22). Equally, the problem of developmental sequences in L2 learning has been addressed by subsequent research that posits a dual route to L2 acquisition in later integrated models (cf. R. Ellis 1994a; Skehan 1998; Towell & Hawkins 1994).⁴⁶ Yet it is perhaps fair to say that the problems in effectively operationalising and testing concepts such as proceduralised and implicit knowledge remain (cf. R. Ellis 2008b).

Despite some of their shortcomings, cognitive approaches to SLA are considered to be an effective method for offering insights into the processes of L2 learning, taking up Selinker's (1972) and Corder's (1967) dictum that an understanding of *how* learners learn a second language is primary. It is also not an *ad hoc* theory, as many of its constructs are taken from cognitive psychology as a whole. Yet compared to UG-based accounts of L2, its explanation of exactly *what* is learned in SLA is weaker, and the precision of its terminology has been seen as a deficit.

1.7 Conclusion

SLA research, some forty years after its inception as an identifiable field of enquiry, is still characterised by facts, opinions, explanations, positions and facts that frequently exist in an uneasy state of complementarity and opposition. SLA is a diverse and divided field of enquiry. Differences are evident in the kind of data researchers collect to investigate acquisition, in the attitudes they hold towards theory development and evaluation, and in how researchers and educators think SLA should inform language pedagogy. (R. Ellis 2008a: 949)

Fourteen years after the publication of the first (1994b) edition of *The Study of Second Language Acquisition*, Rod Ellis draws a similar conclusion about the state of the art of SLA research at the beginning of the 21st century. The field is still marked by an 'uneasy complementarity' of theories, models and accepted 'facts' of the L2 learning process. Ellis adds, however, in contrast to his more sobering (1994b) conclusion, that "what matters most is that SLA has established itself as a vibrant field of

⁴⁶

Indeed, McLaughlin himself admits the possibility of a dual route to L2 acquisition, "[one] highly determined by linguistic constraints, that is predetermined and automatic, and that follows natural acquisition sequences, and [a second] not determined but that requires automatization through controlled processing" (McLaughlin 1987: 149-150).

enquiry with a willingness to explore a wide range of issues by means of alternative paradigms and methods" (2008a: 949).

Yet, as the above overview of changing approaches to SLA research shows, there has been a clear paradigm shift within the field since its inception, namely away from *teaching* and towards *learning*, with a clear focus on the learning process and the role of the learner. The view of the learner as a passive recipient of instruction put forward by behaviourist approaches has changed to one where, within most current thinking, the learner plays an active and creative role in constructing an interlanguage system. The main focus of attention is now finding ways of understanding the processes that take place within the 'black box' of the learner's dynamic restructuring of their knowledge of the L2. Thus approaches are required that focus not on contrastive linguistics but on the psychology of learning, and not on the finished product of learning but on the cognitive processes and strategies that learning involves, and indeed on the general contribution of the individual learner to the learning process.

Many questions as to the nature of the acquisitional processes that the learner undergoes still remain unanswered. The precise role of explicit knowledge is yet to be clarified (cf. R. Ellis 2008a: 900-903), with much research focusing on the possible acquisition-facilitating effects of instruction and consciousness-raising activities (cf. Norris & Ortega 2000). The role of explicit knowledge in the learning process is seen to be far more complex than that imagined by behaviourism, and, indeed, by much research based within a UG framework. Although UG theory has many strengths compared to other approaches such as behaviourism or Krashen's Monitor Model, its suitability for explaining adult L2 learning – particularly in instructional settings where explicit knowledge is seen to play a key role – is limited. Above all, UG offers little insight into actual L2 learning processes, especially at the level of the individual learner.

Thus despite some of the conceptual vagueness of a cognitive approach to adult L2 learning discussed above, it is this information-processing paradigm that offers most potential for explaining the learning processes that learners undergo. It is an approach that provides the possibility to investigate the cognitive elements of L2 learning, particularly with a view to understanding the active role that the learner plays in using explicit knowledge to formulate and test hypotheses regarding the L2 and to use errors in a 'strategic' way. It places the learner at the centre of its focus, yet is also broad enough in scope to account for some of the obvious fundamental differences between L1 and adult L2 learning. In

its modular forms – such as that proposed by R. Ellis (1994a) – it also offers room for the impact of affective factors and motivational elements.

Furthermore, current cognitive approaches to L2 learning provide space for research that is based on micro-theories that intend to explain one component of L2 learning, such as the role of explicit knowledge in an instructional setting as in R. Ellis (1994a) above. Equally, within this framework there is potential for exploratory and interpretative studies based on a more qualitative research paradigm, such as those using thinking aloud data. This can be seen particularly within research into the use of learning strategies in recent studies such as those by Schmidt (2007) or Würffel (2006).

Cognitive approaches to L2 learning have also been supported by technological developments over the last 20 years. Since the 1990s, the benefits of computer technology as a tool both teaching, supporting and understanding the processes of L2 learning have become apparent, with research into computer-assisted language learning (CALL) becoming a discipline in its own right (cf. Chapelle 2001; Kohn 2009; Levy 1997).⁴⁷ The potential of using corpus-based analysis of both native-speaker and L2 learner language, for example, is becoming evident (Braun 2005; Braun, Kohn & Mukherjee 2006), and multimedia-based learning materials are increasingly being seen as a viable method for supporting learning strategies and more effective cognitive processing (Collentine 1998; Duquette 1999; Duquette & Laurier 2000; Rüschoff & Wolff 1999; Tergan 1997; Warschauer 1996).

Thus in view of the strengths it offers for an understanding of L2 learning processes, a cognitive, information-processing framework provides a structure for answering the main questions of this book, namely:

1. How do learners use explicit knowledge of the L2, particularly when learning L2 grammar?
2. What influence do individual learner factors have on learning processes?
3. What facilitating role can computers, in particular multimedia-based materials, have on the learning process?

It is these general questions that will be addressed by this study, with a particular focus on the role and use of explicit knowledge in the learning

⁴⁷ Interestingly, R. Ellis' (2008a) overview of SLA research contains only passing discussion of the role of learner corpora (2008a: 64-65) and no mention at all of CALL as a research discipline.