LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Past and Future in the Time Talk of Teenage Learners of English

A longitudinal study of the acquisition of L2 English
by three students in the language support programme
of an Irish post-primary school

A thesis submitted to the University of Dublin, Trinity College
for the degree of Doctor of Philosophy

Deirdre M. McGarry
School of Education

TRINITY COLLEGE DUBLIN
13 MAY 2015
ACADEMIC REGISTRY

2015

PhD in Edu

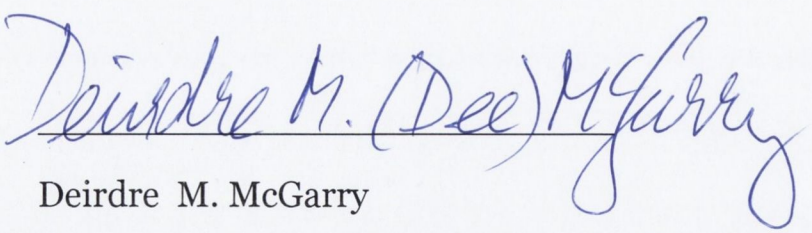
50555779
THESIS
10671

TRINITY COLLEGE
2 9 JUL 2015
LIBRARY DUBLIN

DECLARATION

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.



Deirdre M. McGarry

Thesis 10671

SUMMARY

This study investigates aspects of the English language development of three teenage learners, (native speakers of Polish, Malayalam and Lithuanian) who were enrolled in the language support programme of an Irish post-primary school. The study follows the learners' progress over four school years and adopts a descriptive non-experimental quantitative approach. The analysis is based on samples of written production which were obtained using a process designed to encourage relatively spontaneous writing, with the focus on meaning rather than form. These language samples were transcribed into CHAT format, and then comprehensively coded, providing a learner corpus of over 72,000 words. The data was analysed with the aid of the CLAN suite of programs.

The main focus of this study is the development of past-related and future temporal expression. The learners' data confirms Bardovi- Harlig's findings that the Will Future (*will* +infinitive) emerges early and dominates future time expression, while the Go Future (*be going to)* emerges late (2004b). This study additionally tracks the emergence of the other future forms and finds that the Futurate Present Progressive emerges before *be going to* in the case of two learners and at around the same time in the case of the third. The study also investigates the meanings initially associated with these emerging forms and the way form-meaning associations change in the course of the development of future expression. It finds that, for these three learners, early use of *be going to* is characteristically associated with imminence, rather than with the intention-based meaning which Bardovi-Harlig found to be characteristic of the early use of the form by most of her learners (2004b). After *be going to*'s emergence in this data, *will* is less frequently used to express an imminent future – it becomes more strongly associated with remoteness.

The data supports the order of emergence for the Present Perfect reported by Bardovi-Harlig (2000), with the form emerging in the learners' interlanguage after the Simple Past and the Past Progressive. The data additionally suggests that while the stability of the Simple Past may be a prerequisite, the emergence of the Present Perfect does not necessarily follow soon after that stability is acquired – that the acquisition of the complex semantics of the form may be a lengthy process. The study also examines the spread of past-related forms from prototypical to less typical meaning associations and provides an overview of the order of emergence of form-meaning associations for the temporal system as a whole.

For the analysis of the learners' English language development, the data on temporality was supplemented with data on lexical proficiency and on general accuracy of language use. This analysis includes a longitudinal study of the development of lexical diversity using the D measure.

In this study the analysis of the data provides evidence of the learners' sensitivity to input frequencies and other frequency-related factors, lending support to usage-based models of language acquisition that emphasise the role of input.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr Sean Devitt and Dr Ann Devitt, for their invaluable advice and encouragement; Luisa, Sara and Ana, the three learners who agreed to participate in this study; and finally, and most of all, my husband Aidan for his unfailing patience, understanding and support.

LIST OF ABBREVIATIONS

BNC	British National Corpus
CA	Contrastive Analysis
CHAT	Codes for the Human Analysis of Transcripts
CHILDES	Child Language Data Exchange System
CLAN	Computerized Language Analysis
EA	Error Analysis
ESL	English as a Second Language
L1	First Language
L2	Second Language
LSP	Language Support Programme
SLA	Second Language Acquisition
TA	Tense-Aspect
UG	Universal Grammar

CODES USED IN WRITING SAMPLES

Throughout this thesis, samples of the learners' writing are used for illustration purposes. These include codes which indicate errors, changes and corrections. The relevant codes are explained here. A comprehensive list of the CHAT conventions and the codes used in the transcripts is provided in Appendix One.

[/] This symbol indicates that the preceding word or scoped phrase was deleted and that the deletion was made before the succeeding word was written.
For example: *I maked [/] made a big mistake!*
As soon as the learner wrote *maked,* she realized her error and replaced it with the correct form, then finished the sentence.

[//] This symbol indicates that the change or correction was made after the phrase or sentence was completed.
For example: *I maked [//] made a big mistake!*
The learner only realized her error after she had finished the sentence. She went back, deleted *maked* and inserted the correct form above the line.

[:] Where the learner misspells a word, the correct spelling is given in brackets.
For example*: It was a tragedit [: tragedy].*

[*] Where the learner uses inappropriate marking on a verb, and the resulting form is ambiguous, the target form is given in brackets with an asterisk to indicate that an error was made: *Then I felt [* fell].*

Table of Contents

1 Introduction 1

1.1 Motivation for the study 1

1.2 Research questions 2

2 Generative and Emergentist Accounts of Language Acquisition 7

2.1 Introduction 7

2.2 First language acquisition 9

2.3 Second language acquisition 13

2.4 Linguistic knowledge and SLA research 16

3 SLA Research and the English Verb System 18

3.1 Introduction 18

3.2 Morpheme order studies 22

3.3 Early studies of developmental sequences 26

3.4 Meaning oriented studies of temporality 28

3.5 Form oriented studies of past-related temporality 33

3.5.1 Introduction 33

3.5.2 Order of emergence studies 34

3.5.3 Lexical aspect 38

3.5.4 The role of input: frequency and prototypes 47

3.6 The acquistion of future temporality 57

3.6.1 Introduction 57

3.6.2 Future time reference in English 58

3.6.3 SLA research and future time expression 62

4 Research Design & The Research Context 67

4.1 Introduction 67

4.2 Theoretical framework 68

4.3 Research questions 69

4.4 Research design 70
4.4.1 Introduction 70
4.4.2 A longitudinal perspective 72
4.4.3 Sampling natural language use & implicit knowledge 74
4.4.4 Reasons for choosing written data 77
4.5 The Research & Learning Context 78
4.5.1 Introduction 78
4.5.2 The school 79
4.5.3 The Language Support Programme 79
4.5.4 The writing process 82
4.6 The learners 84
4.6.1 Introduction 84
4.6.2 Learner profiles 85
5 Data Collection, Transcription & Coding 87
5.1 Data collection 87
5.2 Transcription 92
5.2.1 Introduction 92
5.2.2 Adapting CHAT for written data 93
5.2.3 Transcribing the data 95
5.3 Coding 96
5.3.1 Introduction 96
5.3.2 Error coding 96
5.3.3 Morphosyntax analysis 97
5.3.4 Context coding 99
5.4 Analysis 99
5.4.1 Lexical development 100
5.4.2 Accuracy of language use 101
5.4.3 The acquisition of temporal expression 102

6 Aspects of Language Development – Lexical Diversity and Accuracy of Language Use 105
6.1 Introduction 105
6.2 Lexical development 105
6.2.1 Introduction 105
6.2.2 Vocabulary 106
6.2.3 Lexical diversity 110
6.3 Accuracy of language use 115
6.3.1 Introduction 115
6.3.2 Focus on meaning versus focus on form 115
6.3.3 Development of accurate use 119
6.4 Conclusion 126
7 The Acquisition of Past-Related Temporal Expression 128
7.1 Introduction 128
7.2 The Simple Past 129
7.2.1 Past time contexts and the Simple Past 129
7.2.2 The development of appropriate & accurate use of Simple Past 130
7.2.3 Lexical aspect & the learners' use of the Simple Past 136
7.2.4 High-frequency, high-utility verbs 138
7.2.5 The Spread of Simple Past from prototypical to less typical uses 139
7.2.6 Summary 140
7.3 The Past Progressive 141
7.3.1 Introduction 141
7.3.2 The Development of appropriate and accurate use 142
7.3.3 Verb frequencies in Past Progressive 142
7.3.4 Lexical Aspect and the learners use of the Past Progressive form 143
7.3.5 The spread of the progressive to repeated actions and events 144
7.3.6 Summary 146

7.4 The Present Perfect 147
7.4.1 Introduction 147
7.4.2 The emergence of the Present Perfect 147
7.4.3 The learners' use of the Present Perfect 149
7.4.4 Overgeneralization of the Present Perfect 151
7.4.5 Summary 152
7.5 Conclusion 153
8 The Acquisition of Future Expression 155
8.1 Introduction 155
8.2 Future contexts & future forms 155
8.3 The Will Future 159
8.3.1 Introduction 159
8.3.2 Emergence and developing use 159
8.3.3 Summary 163
8.4 Simple Present forms in future contexts. 165
8.4.1 Introduction 165
8.4.2 The Subordinate Future 166
8.4.3 The Futurate Simple Present 166
8.5 The Futurate Present Progressive 167
8.5.1 Introduction 167
8.5.2 Emergence and developing use 167
8.6 The Go-Future 169
8.6.1 Introduction 169
8.6.2 Emergence and developing use 170
8.6.3 Summary 174
8.7 The dominance of the Will-Future & the late emergence of the Go-Future. ... 176
8.8 Conclusion 179

9 Summary and Discussion of Main Findings 181

9.1 Introduction 181

9.2 Lexical development and accuracy of language use 182

9.2.1 Lexical development 182

9.2.2 Accuracy of language use 184

9.2.3 Conclusion 187

9.3 Past-Related Morphology 187

9.3.1 Emergence of the Present Perfect 187

9.3.2 The spread of forms from typical to less typical uses and associations ... 190

9.3.3 Conclusion 195

9.4 Future time expression 197

9.4.1 Order of emergence 197

9.4.2 Form-meaning associations and the development of future time expression 202

9.4.3 Conclusion 206

9.5 An overview of the order of emergence of past and future form-meaning associations 208

9.6 Frequency-related factors and the learners' acquisition of English 210

10 Conclusions 217

10.1 Introduction 217

10.2 Contributions to SLA Research 217

10.3 Proposals for Further Research 220

10.4 Concluding Reflections 222

Appendix 1 Chat conventions and codes used in the transcripts 224

Appendix 2 Transcripts (on attached CD)

References 229

TABLES

Chapter 3

3.1	10 Most Frequent Verbs by Tense-Aspect Category in BNC spoken	52
3.2	10 Most Distinctive Verbs by Tense-Aspect Category in BNC spoken	52

Chapter 6

6.1	Major Word Classes – Types and Tokens	106
6.2	Rank Frequency List of Verbs (by lemma)	107
6.3	Mean *D* Values by period	112
6.4	Changes & Corrections by period	118
6.5	Errors Not Corrected by period	118
6.6	Errors per 100 words by period	119

Chapter 7

7.1	Use of Simple Past periods 1-10	131
7.2	Rank Frequency of Simple Past Verbs	137

Chapter 9

9.1	% of *will* form used to express a near future, before and after *be going to* emerges	205

FIGURES

Chapter 6

6.1 Mean *D* Values by period 112

6.2 Errors per 100 words by period 120

Chapter 7

7.1 Past, Present & Future Contexts in the Data 129

7.2 Appropriate Use of Simple Past periods 1-10 131

Chapter 8

8.1 Forms used in future contexts: % breakdown before the emergence of the Go-Future 157

8.2 Forms used in future contexts: % breakdown after the emergence of the Go-Future 158

8.3 Verbs used with the Go-Future 175

8.4 Will-Future v Go-Future 176

1 Introduction

1.1 Motivation for the study

My personal history as an English language teacher is long and varied. I first worked with learners of English when I was an 18 year old student, volunteering on a language programme for Chilean refugees. Subsequently I spent five years in Africa as a secondary school teacher of English, working with students for whom English was a second, third or even fourth language. In a township school in Harare we had few books, not enough chairs for everyone, often no chalk for the blackboard. Most of the houses had no electricity and students did their homework by paraffin lamps, or sitting under street lights. In Dar es Salaam I worked in an International School, with Tanzanian teenagers there on scholarships, but also with teenagers from privileged backgrounds, whose parents worked for embassies or multinational companies. In Rome I taught English to unemployed refugees and migrants, but also to Italian children and teenagers in a suburban school. Back in Trinity College, I worked on English courses for international students. Since 2004 I have run a large language support programme in a Dublin community school where more than 40 different languages are spoken.

Over the years then, I have worked with many hundreds of people, of different ages and language backgrounds, learning English for different reasons and in different circumstances. As a teacher what has struck me, time and again, is how individual each learner is, and yet how strikingly similar are the patterns that emerge in their learning of the language, the developmental sequences they follow, the milestones that mark the path of their progress towards the acquisition of English.

I have become more and more convinced of the need for language teaching to be informed by acquisition research. Hoping to gain new insights into my own learners' language development, I looked for longitudinal studies of teenage immigrants acquiring English in a formal

education context. However I found that there were few such studies. So I decided to embark on one myself, both to increase my own understanding of L2 English language acquisition, and in the hope that I could contribute something to the body of knowledge in this area. I was particularly interested in investigating whether, and to what extent, L2 language development continued over the entire secondary school period. I chose to focus on temporality because 'the expression of the temporal circumstances and properties of events plays an important role in all linguistic communication' (Dietrich et al., 1995, ix). In addition, I decided to obtain some supplementary information on the learners' general language development using measurements of lexical diversity and accuracy of language use.

1.2 Research questions

To guide this study the following research questions were selected:

Past-related temporal expression: Does the Present Perfect emerge in the interlanguage of all three learners? Does the Present Perfect emerge after the Simple Past and Past Progressive, as reported by Bardovi-Harlig (2000)? If so, does the emergence of the Present Perfect depend on the stability of the Simple Past, as reported by Bardovi-Harlig (2000)?
In the case of the Simple Past, Past Progressive and Present Perfect, is there evidence of continuing development in the use of these forms – of their spread from typical to less typical meaning associations and uses?

Future temporal expression: What future forms emerge in the learners' interlanguages, and what is the order of emergence of these forms? Specifically, does the Will-Future (*will* +infinitive) initially dominate future time reference, with the Go-Future *(be going to*) emerging late, as reported by Bardovi-Harlig (2004b)?
If the Go-Future emerges late, is there evidence in the data to suggest which factor(s) might be responsible for its late emergence?

What meanings are initially associated with emerging future forms? How are form-meaning associations revised following the emergence of new forms?

Lexical diversity and overall accuracy of language use: Does overall accuracy of language use continue to increase? Does lexical diversity continue to develop and does it approach L1 norms during the course of the study?

1.1 Structure of the thesis

Part One Literature Review

Chapter 2 contains some general introductory remarks on issues involved in the study of language. A short overview of Generative and Emergentist accounts of linguistic knowledge and language acquisition is provided in relation to both first languages (2.2) and second languages (2.3). The chapter ends with a comment on implicit linguistic knowledge and the difficulties involved in its measurement (2.4). Chapter 3 looks at second language acquisition research in relation to the English verb system. Section 3.1 provides a brief historical overview of SLA studies from the mid-twentieth century on, leading up to the morpheme order studies (3.2) and the early research into developmental sequences (3.3). Section 3.4 looks at meaning oriented research into temporality, particularly the major study conducted under the auspices of the European Science Foundation (Dietrich et al., 1995). Form oriented research into the acquisition of past temporality is discussed in Section 3.5. First the studies of emergence reported in Bardovi-Harlig (2000) are outlined, then research on the Aspect Hypothesis (Andersen and Shirai 1994) is considered, as is research into the spread of forms from prototypical to less typical uses and associations. The section ends with an overview of studies which have investigated the relationship between input and acquisition. Section 3.6. begins with a discussion of future time

reference in English, before going on to look at Bardovi-Harlig's 2004 studies dealing with the Will-Future and the Go-Future.

Part Two Research Design and Methodology.

Chapter 4 deals with research design and the research context. Section 4.2 briefly outlines the theoretical framework within which the study is conducted and the research questions are presented in Section 4.3. Some issues related to the research design are then discussed in Section 4.4, including the rationale for conducting a longitudinal study, the importance of sampling implicit - as opposed to explicit - linguistic knowledge , and the choice of written rather than oral production data . The research and learning context is discussed in Section 4.5. Information is provided on the school environment and curriculum, on the language support programme to which the students belonged, and on the writing process from which the data was drawn. Profiles of the three learners are presented in Section 4.6.

Chapter 5 looks at data collection, transcription and coding as well as at the programs and procedures used in the analysis. Section 5.1 provides detailed information on data collection, including the contexts in which the written samples were collected and the range of writing tasks involved. In Section 5.2 the reasons for choosing the CHAT format are outlined and there is a brief discussion of the ways in which CHAT was adapted for use with the written data in this study. Coding is discussed in Section 5.3. The section begins by outlining why and how a morphosyntactic analysis of the data was undertaken. The coding of errors and the coding of the tense-aspect contexts created by the learners are also discussed here. Section 5.4 outlines the programs and procedures used for the analyses conducted as part of this study, including the CLAN programs (acquisition of temporality), the vocd program (lexical diversity), the *errors per 100 words* measure (overall accuracy of language use) and the analysis of changes and self-corrections made in the course of writing.

Part Three Analysis and Discussion

While the main focus of the study is an analysis of the learners' acquisition of past and future temporal expression, Chapter 6 provides some supplementary background information on the overall progress made by the learners in relation to vocabulary development and lexical diversity (6.2) and general accuracy of language use (6.3). Areas of significant progress are identified as well as those areas which remain problematic for individual learners right up to the end of the study. Points of particular interest are briefly discussed as they arise, and discussed further in Chapter 9.

Chapter 7 presents a detailed picture of the progress made by the learners in acquiring past-related tense-aspect forms, and relates this to the findings of existing temporality research. Section 7.2 looks at the development of appropriate use of the Simple Past and the spread of the form from prototypical to less typical meanings and associations. Section 7.3 does the same in relation to the Past Progressive. The emergence and developing use of the Present Perfect is discussed in Section 7.4. Points of particular interest are highlighted here and then revisited in the discussion in Chapter 9.

The acquisition of future temporality is the subject of Chapter 8. Section 8.2 discusses where and how future contexts arise in the data, and then provides an overview of the future forms supplied by the learners as well as their relative frequency of occurrence. The developing use of each form is examined in the order it emerges in the learners' interlanguage – the Will Future (8.3) the Subordinate Future Simple Present (8.4) the Futurate Present Progressive (8.5) and the Go-Future (8.6). The dominance of the Will Future and the late emergence of the Go-Future are discussed briefly in Section 8.7 and again in Chapter 9.

In Chapter 9 the main findings of the study are summarized and discussed in relation to the research questions posed at the start. Section 9.2 looks at the development of lexical diversity and accuracy of language use, 9.3 at past-related morphology and 9.4 at future time expression. Section 9.5 provides an overview of the emergence of past and future form-meaning associations. Finally, Section 9.6 looks at the theoretical implications of the study, specifically in relation to the role played by input frequency and frequency-related factors in the learners' acquisition of English.

Chapter 10 outlines the contributions made by the study to the body of second language acquisition knowledge and presents some proposals for future research. The chapter ends with some personal reflections on the study.

Part One LITERATURE REVIEW

2 Generative and Emergentist Accounts of Language Acquisition

2.1 Introduction

The main focus of this study is the development of temporal expression by L2 learners of English and the relevant research is discussed in Chapter 3. The present chapter contains some general introductory remarks on issues involved in the study of language. A short overview of Generative and Emergentist accounts of linguistic knowledge and language acquisition is provided in relation to both first languages (2.2) and second languages (2.3). The chapter ends with a comment on implicit linguistic knowledge and the difficulties involved in its measurement (2.4).

While language, and the ways we use it to communicate meaning, have been the subject of philosophical debate since ancient times - from Confucius and Mozi in China, to Plato and Aristotle in Greece - and while the systematic study of grammar dates back to the Sanskrit scholars of Iron Age India, it is arguably only in the last fifty years that any significant progress has been made in our understanding of what it means to *know* a language and the related issue of how that knowledge is acquired. This progress though significant, is still relatively modest, in the sense that there is much which is not yet established– indeed much which has yet to be investigated.

Some hold that a comprehensive understanding of human language may not in fact be achievable until we are able to look inside the mind of the language user, to examine the conscious and unconscious processes involved. However, despite a century of sustained research, neuroscience has still not developed the tools it needs to record the activity of the individual circuits that underlie a perception or a memory or that give rise to complex behaviours and cognitive functions (Yuste and Church, 2014).

And there are those who argue that even if we could capture, for any given moment, the activity of each of the 100 billion neurons and all of the trillions of connections between them, we might still not have a clear understanding of what was happening in that mind.

The question of how the objective activity of neurons generates a subjective experience has been described as the 'Hard Problem' of consciousness (Chalmers, 1995). While a few (Penrose, 1989; Penrose and Hameroff, 2011) have looked to quantum physics for solutions, others have tried to find the 'neural correlates of consciousness' using brain imaging experiments on language processing (eg Petersen et al., 1988; Raichle, 1998). As the research continues, so too does the debate about whether it is even possible to bridge what Joseph Levine (1983) called 'the explanatory gap'. Chomsky (2002) for example, has suggested that the nature of consciousness may turn out to be a 'mystery' beyond the scope of human cognition, rather than a 'problem' which science can solve.

Whether or not the study of language can shed light on the nature of *consciousness*, there are many who believe it has the potential to provide important insights into the nature of *cognition*. Linguistic behaviour can serve as a 'window' through which to observe the 'entities, processes and structures' of cognition (Dietrich et al., 1995, x) In fact, Evans and Levinson (2009, 447) argue that a central concern of cognitive scientists should be the question of how ' the child's mind can learn and the adult's mind can use, with approximately equal ease' any one of the 'vast range' of different human languages. The question can be put another way : how do we come to use language to delineate time and space, to define reality, assert actuality, to express intention, possibility and volition?

2.2 First language acquisition

At present there are, broadly speaking, two opposing positions on what constitutes linguistic knowledge and how it is acquired. It would be too simplistic to label the debate as one of 'nature versus nurture' though a key difference between the two positions lies in the degree to which the basic structure of human language is regarded as innately determined or as the product of social interaction.

For those in the nativist/generativist tradition, linguistic knowledge is a complex system which cannot be acquired by mere exposure to a language since 'the adult knowledge of language is largely underdetermined by the linguistic data normally available to the child' (Chomsky, 2002, 2). They argue that children can only acquire language as quickly and successfully as they do because human beings are born with a special language faculty, 'a component of the human mind, physically represented in the brain and part of the biological endowment of the species' (Chomsky, 2002, 1). Therefore the child comes equipped with knowledge of phrase structure, and categories like VERB STEM and PAST TENSE MORPHEME are psychological realities for her, not simply descriptive categories employed by linguists, or constructs to be formed on the basis of experience.

Having first introduced the idea of transformational grammar in the 1950s, Chomsky revised and reformulated his theories in succeeding years (e.g. Chomsky,1957; 1965; 1970) until in the 1980s he posited a Universal Grammar with a set of principles which applied to all languages but with individual languages differing from each other because some of these principles contained parameters which could be set in one of a limited number of ways (Chomsky, 1981; 1982). While many generative linguists have adopted this model, there is as yet no general agreement on what constitutes the set of principles and parameters, although some extensive lists have been proposed (e.g. Fodor and Sakas, 2004; Wunderlich, 2004). Over the years generativists have revised or

reformulated theories in the light of first language acquisition research, to accommodate, for example, evidence on the emergence of morphological productivity. Chomsky himself has developed (or revised) his thinking, to give much more importance to the role of lexical learning in language acquisition, and his Minimalist Program (1995; 2000; 2002) now sites parametric variation in the lexicon, principally in the functional (or 'grammar word' and morpheme) categories.

Among generativists there has been considerable discussion about whether functional categories are available from the start, but not evident for pragmatic reasons, or become available at specific maturational stages; or if children build their grammar gradually as they learn the lexicon of their language. This discussion is complicated by the fact that there is no general consensus on the set of functional categories specified by UG (Meisel, 2010, 976).

On the other side of the linguistic knowledge debate are those who believe that language is a complex system whose structures emerge from 'interrelated patterns of experience, social interaction and cognitive processes' (Beckner et al., 2009, 2). Grammar 'comes out of discourse and is shaped by discourse as much as it shapes discourse in an on-going process' (Hopper, 1987, 142). As a result, 'syntax is an emergent phenomenon, not a condition of development' (Ellis, N.C., 1998, 642). Language is learnable from input in part because meanings are directly encoded in surface-level forms and 'no movement, empty, null or silent syntactic elements of any kind are posited' (Boyd and Goldberg, 2009, 418).

For emergentists there is no innate language acquisition device, just the child's need to communicate and her innate capacity to learn – to detect patterns, to make connections, to draw on her experience of the world to solve the problems it presents her with.

Evidence from descriptive linguistics and typology offers insights into the nature of the task faced by the language-learning child. It has also been cited to dispute the existence of a UG, on the grounds that generative theory cannot accommodate the wide variety of human language systems. Although less than 10% of the world's estimated 7,000 surviving languages have been systematically studied, research reveals considerable diversity in sound, morphology, semantic systems and syntax. At one extreme, for example, there are polysynthetic languages with a very high morpheme-word ratio and at the other, isolating languages with relatively few morphemes. There are languages that do not mark tense; there are those that have no fixed word order and those that have no case markings and some that have neither. According to Evans and Levinson (2009) there are languages which rely on dependency relations rather than syntactic constituent structure and there are languages where recursion is a rare or very limited feature of syntax – and, possibly, even completely absent - as claimed by Everett (2005; 2009) in relation to Piraha.

For constructionists, linguistic knowledge is not a system of formal rules and operations but an inventory of constructions (form-meaning pairs) of various levels of abstraction, and various sizes, from morphemes and single words to whole clauses, each of which serves some communicative or socio-pragmatic function. The child acquires this knowledge gradually as she processes many thousands of utterances, noticing correlations between meanings and patterns and contexts

In one account (Tomasello, 2003) she begins by acquiring fixed phrases *(I'm eating it)* from which she then abstracts partially productive lexically specific constructions (*I'm ACTIONing it*) before finally, through analogy, arriving at fully abstract constructions (*the semantic AGENT-ACTION-PATIENT, the syntactic SUBJECT VERB OBJECT*). The child acquires syntactic categories like VERB by grouping together words that perform similar functions and that appear in similar sentence distributions - in

other words through a process of functionally based distributional analysis.

In their examination of an extensive corpus of young children's and mothers' speech, Goldberg et al. (2004) found that there is a strong tendency for certain argument structures (e.g. Verb Locative, Verb Object Object) to be occupied by one single verb with very high frequency in comparison to other verbs used. They suggest that this low variance sample can assist the child in her learning of the construction. According to Goldberg (2006), at a later stage the child will not only be able to employ a range of less typical verbs in each construction, but she will also have realised that some constructions can themselves carry meaning, independently of the particular words used in them. Goldberg argues that the existence of strong correlations between formal patterns and the meanings of the utterances in which those formal patterns appear, facilitates both the child's creative use of language and her understanding of utterances she has never heard before (2006, 34).

Although language behaviour can be described in terms of rules, for emergentists it does not follow that language behaviour is in fact rule-governed. Connectionists, for example, believe that language learning is a consequence of repeated neural network activation that results in stronger and more easily activated connections, in conjunction with certain learning algorithms that adjust the connection strengths between the units in the network in such a way as to decrease the discrepancy between the network's actual output and its desired output (Elman et al., 1998; Mellow and Stanley, 2002) Building on the work of Rumelhart and McClelland (1986), connectionists have used experiments with simple learning mechanisms in artificial neural networks to demonstrate how ruleless subsymbolic associative systems can nevertheless simulate rule-like grammatical behaviour .

Various theoretical models have been proposed for the acquisition, storage and processing of linguistic knowledge, including, for example, the Unified Competition Model (MacWhinney 2005, 2008a) which assigns an important role to the storage of information in lexical maps and the online integration of constructional chunks during language processing. In common with a number of other functionalist accounts, it emphasises the role played by cue availability and reliability in determining the course of acquisition.

2.3 Second language acquisition

Second languages (L2s) are generally acquired under different conditions and in different circumstances to first languages (L1s). Learners are usually older and more cognitively mature, and of course they already possess at least one language. While a large number fail to achieve complete mastery of an L2, many nativists believe that the 'logical problem' of L1 acquisition still applies: that L2 speakers can come to know more than they could reasonably have learned from input alone. This, together with evidence of similar L1 and L2 developmental paths for various morphemes and structures, has encouraged some to argue that Universal Grammar offers the best basis for understanding second language acquisition (e.g. Cook, 2003; White, 2003). However, others working within the UG tradition have argued that second language acquisition is best explained in terms of general learning processes (Bley Vroman, 1989) .

For generativists there are essentially four theoretical positions, which are sometimes described in terms of 'access' to UG and sometimes in terms of the 'states' metaphor which describes the development of the language faculty from a initial state of zero knowledge to a final or steady state of full knowledge. First there is the 'no access' position which hypothesises that UG is not available - possibly because the learner has passed an age-related 'critical period' for natural language acquisition, or because the UG has been 'used up' by L1 acquisition – and therefore the

second language has to be learned by other means, resulting in a grammar which is different, in nature, from that of a native speaker. Secondly there is the 'full access' hypothesis that a copy of UG, unaffected by L1 acquisition, is directly available to the learner enabling her to build an L2 grammar 'from scratch'. In the strong version (e.g. Epstein et al., 1996) the L1 has no influence on the L2. In contrast, the Full Transfer/Full Access hypothesis assigns an important role to the L1 which is seen as the starting point for L2 acquisition, the means by which learners initially access UG, before going on to access UG directly when the L2 does not conform to the L1 settings (Swartz and Sprouse, 1996). The Minimal Trees Hypothesis (Vainikka and Young-Scholten, 1996) asserts that in the case of both L1 and L2 only lexical categories are available at the earliest stages and that functional categories (like agreement) develop gradually in the course of acquisition. Finally there are a number of 'partial access' hypotheses (e.g. Eubank, 1996; Hawkins and Chan, 1997) which agree that second language acquisition is constrained by UG, but differ as to what parameter setting options are available to the L2 learner, for example whether and how they are restricted to their L1 settings.

For Constructionists, linguistic systems are comprised of constructions, therefore language acquisition is the acquisition of constructions and this is as true for L2 as it is for L1. While there are, as already stated, obvious differences between first and second language learners, the overall goal and the specific sub-goals are seen as the same for both - segmenting speech into words and learning the meaning of these words; working out the patterns that govern word combination; interweaving lexical and syntactic knowledge to achieve fluency (MacWhinney, 2008a). As Ellis and Collins (2009) summarise, acquisition is driven by the frequency and frequency distributions of forms, but also by other factors - the salience of the form, its reliability as a predictor of an interpretation and the significance of its functional interpretation. In other words it is not just a matter of how common a form is, but also how noticeable it is, how

obvious its meaning is in the context, and how important understanding that form is to unlocking the meaning of the utterance as a whole.

Salience and contingency issues may help to explain why both L1 and L2 learners find some grammatical features difficult to acquire even though these features occur very frequently in the language the learners are exposed to. They may also help to account for the similarity between English L1 and L2 'orders of acquisition' found in the morpheme studies of the 1970s and 1980s (Bailey, Madden and Krashen, 1974; Brown, 1973; de Villiers and de Villiers, 1973; Dulay and Burt, 1973, 1974; Larsen-Freeman, 1975; Pica, 1983 etc). For example, three of the morphemes in the studies, plural *s*, possessive *s*, and 3rd person singular present *s* are all homophones, and they are each variably represented as the allomorphs *s, z* and *əz,* complicating the association of cue with interpretation. In addition, these bound inflections have low salience in the language stream, being short and low in stress, with a tendency to meld into the words around them. According to N.C. Ellis (2008b) not only L2 learners, but also L1 learners, initially find it difficult to perceive these elements in continuous speech because their language knowledge is not developed enough to provide the top-down support which fluent speakers can rely on.

While the same general learning processes may be involved in both L1 and L2 acquisition, the outcomes or 'end-states' are usually different. Many L2 learners fail to achieve the native-like grammatical competence which is successfully acquired by almost all L1 learners. The limited nature of classroom input and interaction might seem a sufficient explanation in the case of instructed language learning, but even naturalistic L2 acquisition often falls well short, sometimes ending in a simplified 'Basic Variety' of the language, characterised by minimal morphology and pragmatic word order (Klein and Perdue, 1997).

For emergentists second language acquisition involves extra layers of complexity beyond those of L1 (Beckner et al., 2009) . Because L2 learners already have a language firmly entrenched in the mind, cross-linguistic influences can affect the pace at which developmental sequences are traversed; perceptual mechanisms tuned to the L1 can block the learner from perceiving differences in the L2, and conceptual patterns derived from the L1 can shape the way that constructions are put together, leading to non-native categorisations. And, in a general sense, L2 learners know more about how language works than L1 learners do. N.C. Ellis (2008b) for example, describes how pre-existing linguistic knowledge can affect the acquisition of an L2 tense inflection. Children only acquire the meanings of temporal adverbs quite late in development, but second language learners already know how effective such adverbs can be, along with narrative strategies for serialisation, in the communication of temporality. The high salience of these means of expression can lead the L2 learner to pay attention only to them, ignoring the phonologically reduced tense-markings.

Because they are more cognitively mature, L2 learners can treat the second language as an object of explicit learning, of conscious problem solving and deduction (Ellis, N.C., 2003). Some argue that explicit instruction may be necessary for the acquisition of certain constructions, especially bound morphemes. Schmidt (2001, 30) for example, suggests that many features of L2 input 'are likely to be infrequent, non-salient and communicatively redundant' therefore 'intentionally focused attention may be a practical (though not theoretical) necessity for successful language learning'.

2.4 Linguistic knowledge and SLA research

For generative linguists, a dichotomy exists between linguistic competence and performance, and they generally consider samples of actual language use to be both too imperfect a reflection of the underlying system of knowledge and too limited - too finite in the light of the 'infinite

number of interpretable expressions' (Chomsky, 2002, 130) to be suitable for research purposes, preferring to rely on experiments carried out under controlled conditions, using grammaticality judgement tests and tests of metalinguistic knowledge for example. In contrast emergentists believe that the same general cognitive mechanisms underlie both linguistic knowledge and linguistic processing. Language develops through language use and there is no meaningful distinction between competence and performance. Researchers working within this tradition employ a wide variety of methodologies, from computer simulations to the sampling and analysis of natural language use.

However, both generativist and emergentist accounts of second language acquisition broadly agree that linguistic competence consists primarily of *implicit* knowledge – of unconsciously held and rapidly accessed procedural knowledge, as opposed to *explicit* consciously learned knowledge typically accessed through controlled processing - and therefore, for both generativists and emergentists, the goal is to explain how this implicit knowledge is acquired (Ellis, R., 2009).

As many researchers have pointed out (e.g. DeKeyser, 2003; Hulstijn, 1997; Norris and Ortega, 2000), the study of implicit knowledge and implicit learning are areas fraught with difficulties. According to N.C. Ellis, not only is the area beset by measurement problems, but researchers are faced with the dilemma of choosing either the research validity offered by laboratory control and experimentation or the ecological validity afforded by the observation of language learning in its natural environment:

> Every study falls down in one of these respects: consciousness is hard enough to pin down in the laboratory, never mind the classroom. Connectionist models learn language that is a very small sample compared with yours or mine. It is hard to be natural in a loud and claustrophobic fMRI scanner. Real language learning takes tens of thousands of hours, not the minutes of the typical psychology experiment. And so on and so forth, abundantly so.
>
> (Ellis, 2008a, 10-11)

3 SLA Research and the English Verb System

3.1 Introduction

In the mid 20th century SLA studies were dominated by Structuralism (Fries, 1957; Harris, 1951) and Behaviourism (Skinner, 1957). Language learning was largely regarded as a matter of stimulus-response-reinforcement. As Brooks (1964, 49) put it: ' the single paramount fact about language learning is that it concerns, not problem solving, but the formation and performance of habits'. Contrastive Analysis (CA) sought to make this habit formation easier. The target language and the L1 were analysed for similarities and differences in order to predict which features learners would find difficult; the assumption being that 'positive transfer' would occur where structures or features were the same, and 'negative transfer' – interference with habit formation – where they were not (Lado, 1957). Among other things, CA offered teachers information on what errors learners would typically make, and on what structures would need the most classroom attention.

In fact contrastive analyses turned out to be poor at predicting errors: they could not account for many that occurred and conversely those they predicted were often not made by learners (Long and Sato, 1984). Considering the time and effort involved, the return was very small. For these and other practical reasons, and because of its too-close association with Behaviourism, Contrastive Analysis fell from favour in the late Sixties, and the focus shifted to the learner, and the language the learner was actually producing.

With the question of how to teach a language beginning to seem less interesting than the question of how people acquired a language, researchers turned their attention to the internal mental processes that were externalised as L2 oral and written production. The identification of a 'separate linguistic system' (Selinker, 1972, 214) which was 'regular,

systematic, meaningful [...] describable in terms of a set of rules...,' (Corder, 1971, 151) was one of the outcomes of this shift in focus. The learner was no longer regarded as a primarily passive recipient of input, but as someone involved in an interactive creative process. Corder called it an 'idiosyncratic dialect' and Selinker coined the term 'interlanguage'. Its existence came to be generally accepted - a 'third' language, systematic and dynamic, open to change and revision up to the point of fossilisation, involving implicit linguistic knowledge and distinct from both the L1 and the target language.

Much of the work in this area was underpinned by Error Analysis (EA), a discipline in which learner errors were regarded not so much as aberrations or mistakes to be corrected, but as clues to a better understanding of the interactions taking place inside the 'black box' of the learner's mind. Richards (1971) for example, identified three different kinds of errors : interference, intralingual and developmental. Intralingual and developmental errors could not be accounted for by contrastive analysis he said because the former derived from the mutual interference of items within the target language, and the latter derived from the strategies employed by the learner in the course of acquisition. While EA helped to advance SLA research in the early Seventies, it was difficult to resolve issues concerning the identification and classification of errors (Schachter and Celce-Murcia, 1977).

Another issue was EA's assumption that the areas of learner difficulty were revealed by a high frequency of production errors. Schachter (1974) raised the possibility that learners avoid producing constructions which they find difficult both in terms of the actual formation of such structures and the conditions for their use. In support, she presented evidence indicating that Chinese and Japanese learners avoid producing relative clauses in English. Subsequently Kleinmann (1977) provided evidence that Arabic-speaking learners avoid the English passive construction. As Hulstijn and Marchena (1989) note, these two seminal studies encouraged

other researchers to take into account not only learners' overt L2 errors, but also their tendency to avoid L2 forms. EA came to be regarded as limited in a number of important respects:

> Its focus was narrow, since there was a tendency to pay attention only to the errors the learner made rather than to the totality of his output. Furthermore, while it noted the dynamic and developmental nature of the learner's language, the analysis and search for system and explanation were usually confined to one point in time. It was a synchronic analysis. (Devitt, 1992, 132).

While interest in the effect of the L1 on L2 acquisition waned somewhat in the 1970s with the arrival of the morpheme order studies and the search for evidence of language universals (see 3.2), in subsequent decades there has been a notable increase in research into cross linguistic influence (CLI) on diverse aspects of language learning, involving a wide range of research methodologies from the analysis of taped informal interviews to functional magnetic resonance imaging (e.g. Ellis, 2006; Gass and Selinker, 1992; Jarvis, 2000; Master, 1987; Odlin, 1989; Ringbom, 1987; Sagarra and Ellis, 2013; Shirai, 1992; Tolentino and Tokowicz, 2011). While there is still considerable debate as to the nature and degree of L1 influence, recent cognitive approaches to language acquisition research are founded on the 'basic idea that once an L1 is acquired, one cannot process an L2 without the filter of the L1' (Luk and Shirai, 2009, 740).

In the last fifteen years in particular there has been a growing interest in research into multilingualism, and the ways in which the L2 affects L3 - or additional language - acquisition have been the subject of studies in a variety of fields including phonology, morphology, syntax and pragmatics. According to De Angelis and Dewaele (2011, viii): 'The study of CLI has changed radically since it is no longer conceived as a one-language-to-another kind of phenomenon but as a process that concerns all language knowledge in the mind, including the influence of the nonnative languages on the L1.'

The L2 morpheme order studies referred to above, and discussed in the next section (3.2), were inspired by a research project undertaken in the 1960s by Roger Brown and his team at Harvard in the 1960s. This research was to prove hugely influential in both first and second language acquisition research. The study, reported principally in Brown (1973), but also in Cazden (1968) and Bloom (1970), recorded and analysed the spontaneous speech of three English-speaking children, beginning at the point where they had just begun to produce multi-word utterances and continuing for ten months, in the case of Eve, and 20 months for Adam and Sarah.

Brown was interested in both grammatical and semantic knowledge, and the central question addressed by the study was: 'At what point does the child know how to use a given form and when to use it?' (1973, 254). Faced with the problem of establishing criteria for knowledge-acquisition when using spontaneous performance data – because it is not just a question of what the child *is able* to say, but also of what the child *undertakes* to say. – he decided to look at 14 morphemes:

> [G] rammatical morphemes are obligatory in certain contexts, and so one can set an acquisition criterion not simply in terms of output but in terms of output-where-required. Each obligatory context can be regarded as a kind of test item which the child passes by supplying the required morpheme, or fails by supplying none or one which is incorrect. This performance measure, the percentage of morphemes supplied in obligatory contexts, should not be dependent on the topic of conversation or the character of the interaction. (Brown, 1973, 255)

When the data was represented graphically, the gradually rising curve of performance for a child's use of a particular morpheme was observed to be erratic in the earlier stages, finally levelling off within a range between 90 and 100% suppliance in obligatory contexts (SOC). Cazden proposed that the point of acquisition should be defined as 'the first speech sample of three, such that in all three the inflection is supplied in at least 90% of the

contexts in which it is clearly required.' (1968, 435). While acknowledging that any fixing of the point of acquisition must be 'in part [...] arbitrary' (1973, 258) Brown decided to use Cazden's definition - and so subsequently did many SLA researchers.

Brown and his colleagues found that the 14 morphemes were acquired (or approached acquisition) by the three children in an 'approximately invariant' order. Of the nine morphemes from the verb system, the Present Progressive *-ing* was first in the mean order of acquisition, followed by the irregular Simple Past (5th) the uncontractible copula (7th) the regular Simple Past (9th) and the 3rd person singular (10th), with the auxiliaries and the contractible copula coming later. de Villiers and de Villiers (1973) undertook a cross-sectional study of 21 children scored in the same way as Brown's study and using similar criteria of acquisition. They produced an 'accuracy of use' rank order for the morphemes and when Brown did the same for his data the results were strikingly similar. 'It has been made clear that we have a developmental phenomenon of substantial generality,' Brown commented (1973, 274).

For supporters of nativist theories of language acquisition, Brown's study provided an important empirical challenge to Behaviourism, suggesting as it did that acquisition was internally driven rather than environmentally controlled.

3.2 Morpheme order studies

Inspired by Brown's research, Dulay and Burt undertook several studies (1973, 1974,) designed to answer the question: 'is there a common sequence with which children acquiring English as a second language learn certain structures?' Theirs were to be the first of many cross-sectional studies using SOC to investigate second language morpheme acquisition. Instead of recording spontaneous speech, Dulay and Burt analysed data elicited using a proficiency test called the Bilingual Syntax Measure (BSM). Having found evidence for 'a common order of

acquisition' (1973, 256) in their first study of 151 Spanish speaking children, they went on to compare children from two different language groups and concluded that ' the sequence of acquisition of 11 functors [morphemes] obtained for Spanish and Chinese children are virtually the same' (1974, 359). Dulay and Burt claimed that the results provided 'a strong indication that universal cognitive mechanisms are the basis for the child's organisation of a target language, and that it is the L2 system, rather than the L1 system which guides the acquisition process' (1974, 360).

Bailey et al.'s study (1974) involving 73 adults from 12 different language backgrounds, used the same elicitation and scoring methods as Dulay and Burt (1973), and obtained a similar order to that found for the child learners. A similar order was also obtained by studies using other elicitation methods, such as the SLOPE test (Fathman 1975; Krashen et al., 1976). The research suggested that this 'natural order' (Krashen, 1977) was not significantly affected by the age or language background of the learner. Neither was it significantly affected by different conditions of exposure to English, according to Pica (1983) who compared learners receiving formal instruction with those acquiring the language in naturalistic settings.

From the very beginning the morpheme studies were controversial – their data collection, analysis and conclusions. Hatch (1974) for example, criticised the concept of suppliance in obligatory context (SOC) as not taking into account the overuse and overgeneralisation of morphemes (so failing to establish whether learners actually knew the functions of the forms they had acquired; Andersen (1977) said that the scoring procedure resulted in a huge loss of information and Rosansky (1976) questioned the interpretation of rank orders as acquisition orders, arguing that cross sectional studies can only give the state of acquisition at a particular point in time and individual learners' production may vary from month to month.

Nevertheless Larsen Freeman and Long (1991, 92) argued that ' there are too many studies [...] showing sufficiently consistent general findings for the commonalities to be ignored.' In relation to the verb system for example, the same order of acquisition for three morphemes – the progressive *–ing* followed by the regular Simple Past-*ed* and then the 3rd person singular – was reported by many studies, including all twelve of those reviewed in Goldschneider and DeKeyser (2001). However there were inconsistencies – for example where both regular and irregular past were included in the sets of morphemes investigated. As Bardovi- Harlig (2000) points out, while rank order studies (Dulay and Burt, 1974; Larsen Freeman, 1975) found that the regular past preceded the irregular past, hierarchical ordering of morphemes (e.g. Krashen, 1977) produced the opposite result. A longitudinal study carried out around this time (Hakuta, 1976) found that the irregular past came first, and while this finding was based on a single child learner, it was subsequently confirmed by a substantial number of longitudinal studies, including the influential European Science Foundation (ESF) study (Dietrich et al., 1995).

For many critics, the main problem with the morpheme studies was that they focused on the end point of acquisition rather than the developmental process. Andersen, for example, commented: 'Once we became convinced (if we did) that such acquisition orders were real, the question of what this might mean about how learners actually learn second languages was never really answered to anyone's satisfaction' (1991 , 305).

Answers were at least sought by some of those involved in the studies, though their conclusions were tentative. Larsen Freeman, for example, having considered a number of possible determinants of the orders she obtained, suggested that researchers 'should look to the surface form for clues' (1975, 419). Adding that '[a] single explanation seems insufficient to account for the findings', she identified perceptual saliency and

frequency as possible factors, and subsequently (1976) went on to look for correlations between the acquisition order and the frequency of the morphemes in adult native-speaker speech.

Perceptual saliency and frequency were two of five possible determinants investigated by Goldschneider and DeKeyser (2001) in their meta-analysis of 12 morpheme order studies conducted between 1974 and 1993 (the other three were semantic complexity, syntactic category and morphophonological regularity). The results of this meta-analysis must be considered in the light of the criteria for inclusion in the analysis. Of the 25 studies initially identified, 13 were excluded for a variety of reasons. Written data was excluded, for example. Goldschneider and deKeyser also decided to limit their research question to the acquisition of English as a second language, in other words to the acquisition of English in settings where English was the primary language of the learner's environment. As a result a number of studies concerned with the acquisition of English as a foreign language were excluded.

Accuracy results from the oral production data of 924 subjects were pooled and analysed by Goldschneider and DeKeyser. Determining which one was the most important 'causal factor' proved difficult they say, because of the difference in reliability between the predictors, and also the high degree of intercorrelation among them. However Goldschneider and DeKeyser report that a multiple regression analysis found that, taken together, these five factors accounted for 71% of the variance in the acquisition order. They argue that the five factors are in fact aspects of a single underlying one – salience. ' [T]his salience at various levels facilitates the process of induction of grammatical structure from elements of the input' (2001, 37). While the morpheme studies were originally conducted within a nativist theoretical framework, Ellis and Barkhuizen (2005, 77) comment that the results of Goldschneider and DeKeyser's meta-analysis ' are more compatible with current cognitive theories of acquisition that posit a general and minimalist cognitive

structure tuned to notice and process input features [...] rather than a specific and elaborated language acquisition device that directs learners merely to seek triggering evidence from the input'.

Commenting on Goldschneider and DeKeyser's findings in relation to the five factors, Ellis (2006, 173) says : "We must conclude that, to the extent that the order of acquisition of these morphemes is the same in L1 and L2, these factors play a similarly substantial role in first and second language acquisition". However he points out that the studies included in the meta analysis grouped together L2 learners from a variety of different L1 backgrounds, that they did not investigate the ways in which the first language might have a particular effect on the detailed path or rate of second language acquisition.

In recent years researchers interested in cross linguistic influence have revisited the morpheme studies and have found evidence of L1 effects . Luk and Shirai (2009) for example, reviewed studies conducted with native speakers of Chinese, Japanese, Korean and Spanish. They report that while the acquisition order of Spanish L1 learners generally conforms to Krashen's 'natural order', native speakers of the other three languages mostly acquire plural *–s* and articles later than, and possessive *'s* earlier than, is predicted by the 'natural order'.

3.3 Early studies of developmental sequences

As already stated, one of the main criticisms of the morpheme order studies was that they focused on the destination (the point at which a form had been acquired) rather than on the path the learner took to get there. Other researchers working around the same time chose to start from a different perspective and looked at the various changing means by which learners express certain concepts or functions - in other words, the developmental sequences which lead up to acquisition. Among the most notable of these early longitudinal studies are those which investigated the English negative and interrogative systems – for example Milon

(1974), Wode (1976, 1981) who looked at negatives; and Ravem (1968, 1975) and Cancino et al. (1974, 1978) who looked at both systems.

Cancino et al. (1978) undertook a ten month study of the naturalistic acquisition of English by six Spanish speakers – two children, two adolescents and two adults. Based on a frequency analysis of samples of spontaneous speech they identified four main stages for the interrogative. Initially, questions were asked using intonation alone. Then Wh-questions appeared but without inversion. In Stage 3 the copula *be* and the modal *can* were inverted, and subsequently other modals. In Stage 4 inversion began to include the auxiliary *do* and was overgeneralized, with both direct and indirect questions inverted, before finally inversion was confined to direct questions. Similar sequences were found in later studies, including for example Pienemann and Johnston (1987).

Cancino et al. also identified four main stages in the development of the negative, defined in terms of the strategy predominating at the time. At first *no* or *not* were used before the verb; then *don't* was used as a general negative – but it was unanalyzed (unmarked for tense or person). In Stage 3 the modal *can* and auxiliary *be* were negated (*can't, isn't*) and finally at Stage 4 *don't* began to carry tense and person markers (*doesn't, didn't*). Subsequent research largely confirmed this sequence with Wode (1981) adding a fifth stage where the indefinites *some* and *any* are used appropriately.

The early longitudinal studies did not look at these sequences in the context of overall verbal morphology development (although Stauble's 1984 cross sectional study did; and subsequently Pienemann (1998) analysed the data from a number of later studies to produce a set of six stages based on what he considered to be the different levels of processing complexity involved). As Devitt (personal communication) points out, there has been a tendency to treat the negative and interrogative as self-contained systems and to track their development independently of other

features of the language. He argues that there is a good case for a more holistic approach, one that recognises that language is a complex system and that, as Larsen Freeman (1997, 157) puts it: 'the behaviour of the whole emerges out of the interaction of its parts'.

The development sequences found in the early studies were regarded by some as providing strong support for nativist theories of acquisition. Wode (1984) for example, argued that the similarity between child L1 and adult L2 negation sequences could only be explained if a language-specific learning mechanism, rather than general cognition, was involved. Others offered alternative explanations, but what was generally agreed was that the research strongly suggested that the learner's L2 evolved gradually, in stages, towards the norms of the target language. As Selinker summed it up : ' What is nice and clear, and most colleagues now accept this as fact, is that there exist patterns of regularity in the IL (interlanguage) data.' (1984, 338)

3.4 Meaning oriented studies of temporality

According to Bardovi-Harlig (2000), the 1980s saw a shift away from the focus on the acquisition of morphology as form, in favour of a focus on morphology as the surface realisation of an underlying semantic system. She identifies two main strands within SLA research: the first, which she describes as 'meaning-oriented', investigating the expression of concepts through various linguistic devices, and the second 'form oriented' approach investigating the distribution of verbal morphology as an indicator of the underlying semantic system of interlanguage.

Between 1982 and 1987 a major meaning-oriented study was conducted under the auspices of the European Science Foundation (ESF), with the general aim of investigating 'in a systematic fashion, the way in which foreign immigrant workers [..] go about learning the language of their new social environment' (Dietrich et al., 1995, 2). The study was longitudinal and cross-linguistic, covering a period of 30 months and involving five

target languages. As it was designed, in part, to test for L1 influence, the 40 learners were drawn from six different language backgrounds, making it possible to study a range of L1/L2 combinations. Recordings were made at monthly intervals, usually of conversations and personal narratives, but also role plays and film retelling tasks.

Temporality was one of the main research topics chosen ' since the expression of the temporal circumstances and properties of events plays an important role in all linguistic communication' (Dietrich et al., 1995, ix). 21 of the learners were involved in this particular research, including four learners of English – two Punjabi speakers and two Italians. The principal finding of the study was that the acquisition of temporal expression is continuous and gradual, with learners, irrespective of first language or target language, passing through three stages - pragmatic, lexical and finally morphological (Klein et al., 1995).

In the first stage, the learners' utterances are 'heavily context dependent' and characterised by the use of lexical items - mainly bare nouns, some adverbs, adjectives and nominalised verbs, - with no clear sign of grammatical organisation. For the expression of temporality they rely on adverbs or adverb-like expressions and the discourse strategy of narrating events in chronological order. Conversations require considerable 'scaffolding' – often the learner cannot make themselves understood without the co-operation and support of their interlocutor.

In the second stage, there are still no finite constructions, except for some rote forms. Utterances typically consist of uninflected verbs and their arguments, and, in contrast to stage 1, there is evidence of some clear organisational principles at work in the way words are put together. The learners' repertoire of temporal adverbials increases steadily to include a range of anaphoric and deitic adverbials (e.g. *then, after; now, yesterday*) as well as adverbials of frequency (e.g. *one time, always, often*) and duration (*two hour, four day*). In addition there are now some boundary

markers – words (usually verbs) like *start, finish* which allow learners to mark the beginning or end of a situation.

Klein et al. call this second, lexical, stage the 'Basic Variety' of the L2 and say that while its temporal system is very simple, with clear shortcomings that effect communicative efficiency, it is also very versatile. ' In principle, it allows an easy expression of when what happens or is the case, provided that there are enough adverbials and that it is cleverly managed' (1995, 269). One group of learners in the ESF study never went beyond this stage but steadily improved their expressive power in the L2 by extending and enriching their vocabulary, (particularly adverbials) and perfecting their handling of the system. As Bardovi- Harlig (2000) comments, previous SLA research had already noted the role played by adverbials, but the ESF study and other meaning-oriented research (e.g. Meisel 1987; Schumann 1987) served to clarify just how central that role is.

While the basic variety may be versatile, Klein et al. point out that it 'strongly deviates from the language of the social environment [...] and stigmatises the learner as an outsider' (1995, 272) and they observe that, for those of their learners who moved on to the morphological stage of development, this social reason was a more important motivating factor than ease of communication.

The structure and function of the basic variety was found to be more or less the same for all learners and Klein et al. suggest that it could plausibly be said to reflect more or less universal properties of language. For further development to take place, however, the learner has to adopt 'the peculiarities of the language to be learned' (1995, 269) and for that reason they say they found it more difficult to identify general properties for this third stage. A number of common features were, however, observed in the ESF study - and confirmed by other studies (e.g. Bardovi- Harlig, 2000) . First, in the case of both temporal adverbs and morphology, development

is slow and gradual 'with no distinct and sharp developmental steps'. Secondly, form precedes function - appropriate and inappropriate usage co-exist for a long time before full control of the functions of a form is achieved. Thirdly, tense marking precedes aspect marking: 'it is true that learners of English may have perfect forms and, especially, progressive forms at an early stage, but in no case do we observe an early functional use of these forms' (1995, 270). Fourthly irregular morphology precedes regular morphology. Klein et al. say that this is the case despite the fact that past tense formation is very simple for the regular forms but often 'a nightmare' for the irregular past. Pointing out that irregular verbs are typically frequent and that a regular ending such as the English *–ed* may lack perceptual salience, they suggest that 'second language acquisition as observed here is inductive and heavily input oriented' (1995, 271).

Lavinia, an Italian speaker, was the only one of the study's learners of English to reach the stage of morphological development. However her progress was generally in line with developmental sequences reported in other studies of L2 English. The past irregular, for example, was routinely being used appropriately before the first occurrence of a regular past form at month 13. In month 16 there was increased use of the past regular and in month 21 the Present Perfect was being regularly used to express aspect, and the Past Perfect made its first appearance.

In relation to adverbials, which were shown to do a considerable amount of the temporal reference work in the lexical stage of development, Klein notes that Lavinia's repertoire 'is slowly enriched and refined' (1995b, 47) but does not comment on the extent of her usage. Bardovi- Harlig (1992) however found evidence that learners' use of time adverbials decreases as use of past-time verb morphology increases. In a subsequent 1994 study she found that as the use of past tense stabilizes, and learners start to make reference to anterior (past in past) events, they begin the cycle of dependency again, using adverbials to mark deviation from chronological order, both before, but also after, past perfect morphology emerges. 'It is

possible' she comments 'that this cycle of lexical to morphological marking occurs throughout the tense-aspect system whenever new forms – and meanings – are added to the system' (2000, 47).

Bardovi-Harlig suggests that one factor contributing to the functional load borne by adverbials might be the difficulty that learners may experience in comprehending verbal morphology. According to Van Patten (1996) for example, learners process for meaning before form and they prefer to process lexical items over grammatical items for semantic information. Bardovi-Harlig cites evidence from experimentally-based processing studies which appears to show that lexical cues are more important to learners than morphological cues. Lee et al. (1997) for example, found that when learners of Spanish heard a passage containing adverbials they reconstructed significantly more of the related events using past reference than did the learners who heard a version of the passage without adverbs. And a study by Boatwright (1999) using reaction times rather than recall scores yielded the same results. Bardovi-Harlig comments that in texts without adverbials learners may try to use context rather than verb forms to construct past reference. 'In fact the comprehension strategies of some learners may focus their attention away from processing the information carried by the verbal morphology, even though processing the verbal morphology is necessary in the acquisition of the form-meaning association' (2000, 44).

The final finding reported by the ESF study relates to cross linguistic influence. Klein et al. (1995) did find 'some transfer phenomena' - in a few cases, for example, the Basic Variety base verb form varied depending on L1, and learners occasionally borrowed words from their first language. However the researchers said that what was much more striking was the lack of L1 influence where they would have expected it – for example in relation to the presence or absence of aspect marking in an L1. 'We must conclude therefore, that there is no significant [L1] influence in the acquisition of temporality' (1995, 278).

A significant contribution made to SLA research by the ESF study, and similar meaning oriented analyses, was that they rejected what Klein calls 'the inflexional paradigm bias' (1995a, 18) and looked not only at tense and aspect marking, but also at the many other ways in which temporality is encoded in natural language. And the ESF study broadened the scope of such research further by encompassing within its view not just the development of temporality, but the development of the interlanguage system as a whole.

3.5 Form oriented studies of past-related temporality

3.5.1 Introduction

Taking as its starting point a semantic concept, a meaning oriented (or function-form) analysis can reveal all the various morphological forms used to express that concept. But in order to achieve a fuller understanding of acquisition processes it is also necessary to look at how individual forms develop, and at how they are deployed in different interlanguage contexts before finally becoming established in their own appropriate environments. For this, another kind of analysis is needed – one which follows the form from first appearance to, or towards, acquisition.

Such form-oriented (form-function) analyses differ from the earlier morpheme studies in that they focus on *emergence* rather than *acquisition* and on *sequences* rather than *orders*. 'Thinking in terms of sequences in studying the acquisition of tense and aspect emphasizes that a single grammatical subsystem is being investigated and that the parts of the systems are intrinsically related to the other parts and will have an impact on each other as the system develops' (Bardovi-Harlig, 2000, 95).

Acquiring the tense-aspect system for the English past, for example, involves not only the task of learning how to construct the different forms

accurately – the Simple Past *-ed,* Past Progressive *was/were + V-ing,* Present Perfect *have + V-en,* and Past Perfect *had + V-en* – but also the much more difficult task of sorting out the contrasting meanings and uses of these forms. As a result, when a new morpheme enters the interlanguage system it only gradually comes to be used appropriately, as its form-meaning associations are established in relation to the form-meaning associations of other morphemes. As the ESF studies established, form precedes function (Dietrich et al., 1995).

3.5.2 Order of emergence studies

This process is examined by Bardovi-Harlig (2000) in a major form-oriented study of the emergence of past morphology in English. The study sheds light on how learners come to distinguish the meaning and use of each of the tense-aspect forms from those of their 'semantically close neighbours' and it establishes an order of emergence. 1576 written texts, mainly journal entries, and 175 oral texts, mostly interviews, were collected over a two year period from 16 college students enrolled in an intensive English language programme. The mean length of observation was 11.5 months, ranging from 7 to 17.5 months.

Bardovi-Harlig counted types rather than tokens of verbs; each verb form (e.g. *got, made*) was counted only once in each sampling period. She used type analysis to avoid inflation of the rates of appropriate use by multiple occurrences of common verbs like *was* and *went.* 'Thus a type analysis provides a conservative view of the acquisition of tense-aspect morphology' (2000, 139). Emergence of a tense-aspect form was defined as the appropriate use of the form with three distinct lexical verbs. And because 'time is a poor predictor of development' the development of the past tense was used as a ' system-specific means of gauging development and comparing learners' (2000, 142). A form was coded as Past Progressive only if the *–ing* morpheme was present; The Past Perfect (or Pluperfect, the term preferred by Bardovi-Harlig) was identified by the

use of *had* with a main verb, and the Present Perfect by the use of *have* or *has* with a main verb (or, in the case of the Perfect Progressive, *-ing* on the main verb or *be* plus the main verb).

Although the learners were assessed to be at approximately the same level of proficiency at the start of the study, they displayed individual differences in rates of attainment, and eventual proficiency. Nonetheless a clear order of emergence was evident: the Simple Past, followed by the Past Progressive, then the Present Perfect, the Present Perfect Progressive, and finally the Past Perfect. As Bardovi-Harlig notes, the order is most easily observed in the language samples of learners whose verbal morphology was less developed at the outset of the study. Three of the most advanced learners, for example, showed emergence of all three forms within two consecutive 15 day sampling periods.

While it was not specifically designed to test for the influence of L1 on L2 acquisition, the study found no obvious evidence of it. Four L1 groups were included - Arabic, Japanese, Korean and Spanish – and '[none] seems to have a unique profile regarding either number of forms produced or emergent form-meaning associations en route to the target associations'(2000, 182).

Past Progressive

According to Bardovi-Harlig the emergence of the Past Progressive does not seem to be dependent on the stability of the Simple Past but only on the prior emergence of the Present Progressive, which itself follows the emergence of the bare progressive . Half of the learners were recorded as first using the Past Progressive at a stage where their appropriate use of the Simple Past was below 80%. She points out that the 80%-90% level is often considered to indicate acquisition and suggests that these results show that the two forms, Simple Past and Past Progressive, are developing at the same time. In the case of first use of the Present Perfect there was a mean group rate of 86% appropriate use of the Simple Past. For the

Past Perfect the rate was 89%. This suggests that Simple Past use needs to have stabilized before these two forms emerge.

Past Perfect

Bardovi-Harlig says that the Past Perfect is ' the undisputed latest starter' (2000, 169). While the Past Progressive and Present Perfect emerged for all learners in the course of the study, the Past Perfect emerged for only half of them. There were 311 instances of the Past Progressive in the written corpus, and 457 instances of the Present Perfect, however the Past Perfect was only used 192 times, and 74 of those uses were by one learner, Carlos. Learners also produced many more ill-formed instances of the Past Perfect than was the case with the Present Perfect.

Present Perfect

93% of the all the Present Perfects in the written corpus were properly formed, and over 80% were used in an appropriate Present Perfect context in both present tense texts and past tense ones. 64% of the Present Perfects occurred with no adverbial, and the rest occurred with a wide range of adverbials of which the most frequent were *for* and *since* (although *never, already, ever, yet* and *just* were also found, along with adverbs of frequency, such as *everyday*, to indicate the iterative (repeated) nature of events sometimes described by the Present Perfect. The Present Perfect Progressive emerged later and was used by 12 of the learners, 45 times in total. A very high percentage of this use was appropriate – 89% - which Bardovi-Harlig says reflects the fact that the appropriate use of the Present Perfect is established before the progressive form appears.

Bardovi-Harlig emphasises that while her study found a clear order of emergence this does not suggest that acquisition of a form is in any sense finished before the interlanguage system admits the next one : 'in the acquisition of verbal morphology form-meaning associations cannot be said to be truly complete for any given tense-aspect form until the entire system is complete' (2000, 175) This is because acquiring the meanings

and uses of a new form involves a reassessment and re-mapping of the meanings and uses of other forms. This process can be observed in the ways emerging forms are sometimes overgeneralized , invading the environment of an existing form, resulting in a temporary decrease in the appropriate use of that form. In the case of one of Bardovi-Harlig's learners, Toshihiro, for example, his overuse of the newly emerging Present Perfect resulted in his rate of appropriate Simple Past use falling from 100% at month 1 to 85% at month 1.5.

Overgeneralisations

There were 65 identifiable overgeneralisations of the Present Perfect in the data, of which 41 occurred where the simple past would be preferred by native speakers; a further 15 were used instead of the Past Perfect and 7 instead of the present tense. Bardovi-Harlig suggests that the reason learners show the greatest overuse of the Present Perfect in Simple Past contexts may be due to the fact that the degree of semantic overlap is strongest for this pair. She points out that the two forms uniquely share truth value – the Present Perfect sentence is true in all cases where the Simple Past sentence is true: *if Max has seen the president* then *Max saw the president*, and vice versa.

Overgeneralisation appears to characterise early use of the Past Perfect. Even Carlos the most advanced learner in the study and the most accurate in his use of the form (82%), used the Past Perfect inappropriately in five of his first six attempts. Overall almost one third of the Past Perfect forms in the corpus were used in contexts where other forms would be preferred (over 42% if Carlos is excluded) and the overgeneralisations are split almost equally between Simple Past and Present Perfect contexts. Shared semantic features could explain the association between Past Perfect and Simple Past, but not that between Past Perfect and Present Perfect. Here Bardovi-Harlig suggests that the formal similarity between the two perfects is likely to be at least partly responsible for the learner's overuse of one form in a context which calls for the other.

Factors affecting the order of emergence

In relation to the order of emergence , Bardovi-Harlig notes that it is unlikely any one factor could provide a satisfactory explanation. Morphosyntactic complexity, for example, could be part of the reason why the Present Perfect is acquired later than the simple past, but it would not explain why the Past Perfect comes after the Present Perfect since their forms are equally complex, differing only in the tense marker. However, unlike the Present Perfect, the Past Perfect is never essential in its environment– its function can be performed by the simple past and adverbials. Therefore functional load could be a factor in the relative ordering of the two forms. Bardovi-Harlig suggests that input frequency is another factor likely to contribute to determining the acquisition order, and it is a factor which has received increasing attention in recent years from SLA researchers working within an emergentist framework (e.g. Boyd and Goldberg, 2009; Bybee, 2010; Ellis et al., 2013; Wulff et al., 2009).

3.5.3 Lexical aspect

Bardovi-Harlig's study investigated the order in which different morphemes emerge and establish themselves with respect to other morphemes in the tense-aspect system. Form-oriented studies have also examined the way an emerging morpheme spreads, from early use with a few verbs, to productive use with many. This body of research - mainly cross-sectional and experimental in design, but also longitudinal - has looked at the influence of lexical semantics and the role of discourse structure, and more generally at how input effects such as frequency and saliency affect the path of acquisition.

In the 1970s language researchers first noticed that children's use of morphology appeared to be affected by lexical aspect - in other words by the inherent temporal characteristics of verbs (Antinucci and Miller, 1976; Bloom et al., 1980). A decade later the issue was taken up by Andersen and other second language acquisition researchers, leading, after some

revisions, to the formulation of the hypothesis that 'First and second language learners will initially be influenced by the inherent semantic aspect of verbs or predicates in the acquisition of tense and aspect markers associated with / or affixed to these verbs' (Andersen and Shirai, 1994, 133).

Applied to English this suggests that the first verbs to attract past tense marking will be those whose inherent meaning implies an end point ('achievement' verbs like *find, arrive*; and 'accomplishment' verbs like *build, give*). In a similar way, 'activity' verbs like *play, walk, look*, which have the inherent meaning of continuing over time without a specific end point, will be the first found with progressive –ing. And when a morpheme begins to be used appropriately – in its proper tense-aspect contexts – then the same pattern will apply, with, for example, appropriate use of the Simple Past appearing initially with telic (or end-point) verbs before spreading to atelic activity verbs and then to state verbs.

So in the early stages of language learning what appears to be random variation in the use of morphology could in fact be the result of connections made between the meaning of the verb and the meaning of the grammatical marker. The apparent randomness may, to some extent anyway, represent a spread from prototypical to less typical.

According to Ayoun and Salaberry (2008) the Aspect Hypothesis has generated the largest body of empirical studies on the L2 acquisition of tense-aspect marking, either directly or indirectly (with studies testing other hypotheses also testing the prediction of the Aspect Hypothesis). A range of data collection methods and analyses have been employed. As Bardovi-Harlig (2000) has pointed out, some studies adopt an across-category analysis, investigating where various morphemes occur, and others adopt a within-category analysis investigating how the different lexical aspectual categories are marked by the learners. While there have

been some longitudinal studies, cross-sectional studies predominate, many of them using elicited narratives or cloze passages or a combination of both. In relation to elicitation methods, a number of issues have been identified and discussed in the literature. Certain types of predicate occur more frequently than others in natural language use for example, and while some within-category analyses have chosen to use cloze passages where the aspectual classes were balanced exactly or approximately (e.g. Bardovi-Harlig and Reynolds, 1995; Collins, 2002) others have not – citing, for example, their preference for using an authentic text (e.g. Ayoun and Salaberry, 2008). On the other hand, where personal and impersonal narratives are elicited or collected, there may be considerable variation in the number of tokens produced by individual learners. Longitudinal studies have tended to use mainly conversational interviews or recordings of spontaneous speech (e.g. Housen, 2002a ; Lee, 2001; and Rohde, 1996 for English L2; Andersen, 1991 for Spanish L2; Giacalone Ramat and Banfi, 1990 for Italian L2).

Both cross-sectional and longitudinal studies have reported support for the Aspect Hypothesis in relation to the English past tense. Rohde (1996) for example, analysed data from a 5 month longitudinal study of two German speaking children (aged 6 and 9) and found a distributional bias for both regular and irregular past inflection Counting types rather than tokens, most of the verbs inflected for past tense were achievements, with regular past showing a stronger bias (56/70 or 80%) than irregular (52/84 or 62%). In a 13 month study of two adolescent Korean speakers Lee (2001) found that both learners marked past tense first on event verbs (achievements and accomplishments). The analysis also revealed the development of lexical aspect marking, from event verbs to activities and states. Lee noted that these young learners' use of the Simple Past was dominated by irregular verbs, with regular past forms only appearing in the later stages of the study, and with the younger learner rarely producing the regular form (the dominance of irregular verbs in past contexts was also noted by other studies including, for example, Housen

(2002a, 2002b) and Salaberry (2000)). Another longitudinal study, Rocca (2002), followed three young Italian children (aged 7 and 8) over their first six months of learning English in the UK. She reported that past morphology was initially restricted to telic (achievement and accomplishment) verbs and then gradually extended to activity verbs and, to a lesser extent, state verbs. She also reported that both irregular and regular past tense forms were strongly associated with telic predicates.

Chan et al (2012) performed a quantitative analysis on the L2 English data from the ESF corpus (Perdue 1993). They looked at the frequency of occurrence of past tense and progressive markers in the developing language of four adult naturalistic learners, two native speakers of Italian and two Punjabi speakers. Copula and auxiliary *be* were excluded from this across-category analysis, leaving a total of 290 tokens inflected for past morphology. Two of the learners produced only 17 tokens between them, but the other two learners' data showed a strong association of past marking with achievement verbs. 73% (159/218) of Lavinia's past-marked verbs were achievements, and 100% (55/55) of Ravinder's were.

For their cross-sectional study, Bardovi-Harlig and Reynolds (1995) devised a 62 item cloze test sampling all four lexical aspectual classes, and administered it to 182 adult learners at six different levels of proficiency. They found that event verbs showed higher use of past than activity or state verbs. Collins (2002) got essentially the same results when she used the same cloze test (and subsequently a revised test) with French speaking adults. She also found that her learners continued to be influenced by lexical aspect even as their use of past tense morphology became more productive. Based on an analysis of oral and written narratives from 37 adult learners, Bardovi Harlig (1998) also supported the hypothesis in relation to the Simple Past – a pattern of high use with event verbs and lower use with activities , even for the more advanced learners. In the oral data there was a very clear progression of past tense use from achievements to accomplishments to activities.

In her 2002 study, Collins noted that once her French speaking learners were able to use the Simple Past appropriately 50% of the time, there was a marked increase in overgeneralization of the Present Perfect to contexts where the Simple Past was required. Collins pointed out that the Present Perfect is similar in form to, but different in function from the French *passé compose* and that the overgeneralization of Present Perfect was not observed in Bardovi-Harlig and Reynolds (1995), a study which did not include learners with L1 French. Collins took a closer look at L1 influence in her 2004 study on the acquisition of past temporal expression by 70 Francophone and 69 Japanese speaking college students. The learners, who were assessed to be at 'upper beginner' or 'intermediate' level, completed a cloze test consisting of 25 short passages. For the analyses, learners were divided into six groups based on their knowledge of Simple Past morphology, as measured by their overall appropriate use of past in the cloze test. For both Ll groups, the marking of past spread from achievements and accomplishments through activities to states, and the preferred non-past responses for states and activities were present and progressive respectively. The Japanese learners (at 81% appropriacy) were more successful at using the past than the Francophones (74%), a difference which Collins attributed to the Francophone learners' overuse of Present Perfect with achievements. When non-past morphology was used by the less proficient learners in the study, perfect forms were the most frequently supplied by both L1s, however the Francophone learners did so most frequently with achievements. Collins concluded that the magnitude, rather than the direction, of the lexical aspect effect with achievements was the only significant Ll difference found by the study.

In a cross sectional study involving 21 French speaking young adults who had been learning English in school for an average of just under 8 years, Ayoun and Salaberry (2008) also investigated whether the learners overused the Present Perfect in Simple Past contexts due to its morphological similarity with the *passé composé*. They found that there

was an L1 effect in Present Perfect use, albeit limited to some learners, and that the Present Perfect was the alternative form used most often with event verbs, especially with increasing levels of proficiency (although it remained relatively low). Ayoun and Salaberry pointed out that both they and Collins used intact groups of learners and that it is possible that 'the individual differences that show up in the results of these three studies reflect the fact that there might be a more fine-grained distribution over proficiency levels that could be revealed with the use of more meticulous procedures to classify proficiency levels' (2008, 583).

Overall, in relation to the Aspect Hypothesis, Ayoun and Salaberry (2008) reported that the results of the two elicitation tasks they used suggest that lexical aspect is a strong predictor of the use of past tense markings. However the study's support for the Aspect Hypothesis was qualified, they said, because the data from the narrative task revealed that state verbs were marked with past tense morphology more consistently than event verbs (accomplishments and achievements). Ayoun and Salaberry speculated that it is possible that learners were reacting to a distributional bias in input data that suggested to them that event verbs are consistently marked with both Simple Past and progressive, whereas states tend to be marked consistently with Simple Past only. 'Thus, in line with Andersen's (1989) one-to-one and multifunctional principles, we expect an unbalanced distribution of past tense markers in the L2 data by which the verb types with the most consistent marking of past tense will be incorporated into the developmental system first' (2008, 580). Ayoun and Salaberry also speculated that the distinct outcome in the narrative task might be associated with task effects , that a cloze task might allow for the monitoring of language form whereas narratives may tap more directly into the learners' competence .

A number of other studies reported findings in relation to state verbs which seemed to be contrary to the Aspect Hypothesis. In a longitudinal study of four German speaking children acquiring English in a naturalistic

setting, Rohde (2002) found a very high past marking rate for state verbs (80-100%). In fact all four learners were more likely to use past marking on state verbs than on verbs from any other aspectual class (however it should be noted that Rohde included the verb *be* in the analysis and that in the case of two learners there was a small number of verb tokens overall, and less than 14 state verb tokens between them). Rohde observed that while this finding was not compatible with the Aspect Hypothesis, another finding was - event verbs were much more frequent in past contexts in the data than states and activities. Rohde went on to discuss a set of factors (e.g L1 influence, input) which might influence the learners' use of verb morphology and then suggested that it might be more appropriate to talk in terms of ' 'an aspectual effect' which can vary in strength rather than of an Aspect Hypothesis that is either supported or not' (2002, 216).

Housen (2002b) analysed oral data (conversational interviews and guided speech tasks) from 46 students learning L2 English at international schools in Belgium. The students, who had Dutch or French L1s, ranged in age from 9 to 17. The analysis of the data for the two groups of learners in the earlier stages of language acquisition showed that past tense marking was found first on frequently occurring state verbs (e.g. *had, saw*). Regular (–ed) morphology did appear first with achievements, before spreading to accomplishments, then activities and states. However, Housen reported that it was not possible to state with certainty that the learners' use of past on regular verbs followed the pattern of development predicted by the Aspect Hypothesis, because the token frequencies for regular past were too low.

Housen (2002a) reported similar results for a longitudinal study of the acquisition of L2 English by a young Dutch speaking child. This study found that irregular past was equally likely to be found with states as with event verbs; the number of regular forms in the data was too low, he said, for any conclusion to be drawn in relation to the development of regular

past morphology. Drawing on the 'dual mechanism' processing theories proposed by Pinker and his colleagues (e.g. Pinker and Prince, 1994; Pinker, 1999) Housen suggested that the Aspect Hypothesis might be valid for regular past only, assuming the acquisition and use of regular past mainly involved productive, symbol manipulating rule-learning, while the acquisition of irregular past mainly relied on associative memory: '...one could speculate that conceptual-semantic notions (prototypes) such as stativity, durativity and telicity play a steering role in the process of morphological rule-learning, which mainly affects regular morphology [...] but not or less so in associative learning, which mainly affects irregular forms such as went, go' (2002a, 188).

Reviewing Housen's 2000a study, Shirai (2010) pointed out that a type count analysis of Housen's data showed that irregular morphology was in fact more strongly associated with telicity than regular morphology was, except in the first two time periods (of six). Shirai argues that, while L2 research to date has been inconclusive, the available empirical evidence is more consistent with a single mechanism model for past tense acquisition – in other words with regular and irregular forms being acquired through the same cognitive processes – rather than the dual mechanism model.

Among recent lexical aspect studies, Chan et al. (2012) specifically investigated the evidence for a dissociation between regular and irregular past marking. They compared both type and token counts for irregular and regular past forms in the L2 English data of the ESF corpus and found that the two forms have similar distribution in achievement verbs for each learner. In fact, Lavinia (who produced most of the past-marked tokens in the data) exhibited similar token frequency distribution between regular and irregular verbs across lexical aspectual classes. Chan et al. conclude that there is no sign of regular/irregular dissociation in the data and so their study does not support a dual mechanism model for the acquistition of past morphology.

In relation to the English progressive, support for the Aspect Hypothesis has also been reported by Housen, Collins, Lee and Bardovi-Harlig, among others. Lee (2001) found that progressive marking was first encoded on activity verbs, as the hypothesis predicts. The spread of progressive marking to event verbs, however, was not observed for the two learners in her longitudinal study as their use of –ing with accomplishment and achievement predicates did not rise above 20% throughout the data collection period. Housen (2002b) found that in the earlier stages of acquisition the use of the –ing marker was mainly restricted to activity verbs, and while it then spread to other verbs as it began to be used appropriately, there was still some evidence of a distributional bias towards dynamic, durative verbs (activities and accomplishments) – a bias he also observed in the native speaker data. Collins (2002) found that the Past Progressive was more readily used with activity verbs than with achievement and accomplishment verbs and Bardovi- Harlig (1998) found that the progressive (with or without the auxiliary) was used to a much greater extent with activities than any other aspectual class.

Bardovi- Harlig (1998) also looked at the influence of discourse structure on the learners' use of aspect. She was specifically interested in the distinction between foreground and background which is believed to be a universal of narrative discourse (Dahl, 1984) with a speaker 'foregrounding' the story line, the chronological sequence of main events, and consigning everything else – explanation, evaluation, interpretation, prediction – to the background. According to Andersen and Shirai (1994, 152) there is a strong association between aspectual class and discourse function; achievements and accomplishments typically perform the central role of laying out events and therefore are 'logical recipients' of Simple Past marking, while activities typically serve a supporting role and tend to be marked for the progressive. Nevertheless because foreground actions are presented and interpreted as though they were punctual, complete and sequential (Reinhart, 1984) even an atelic apunctual

activity once it is used as a foreground predicate takes on the features of punctuality and completeness, attracting the simple past marking.

Bardovi-Harlig found that (for learners with limited linguistic resources) the effect of grounding is especially strong for accomplishments and activities, which both showed higher rates of Simple Past inflection in her data when they were foregrounded. She also found that use of the progressive with activities was limited to the background. Achievement predicates did not seem to be affected by grounding she says – they were the most likely to be inflected for Simple Past regardless of the role they were playing in the narrative. (This is in line with the fact that achievements, being both telic and punctual are closest to the prototypical meaning of 'pastness'). Bardovi-Harlig speculated that using verbs to construct discourse may be one way in which learners come to expand their interlanguage prototypes and move towards the kind of creative command of tense-aspect which native speakers have. When learners foreground activities, she said, 'they may be gently nudged into marking activities with the past by the universal pressure to distinguish the foreground from the background. Once a learner's interlanguage accommodates foreground activities with simple past, this may lead to other uses of past activities; similarly, progressive background accomplishments may open the way for other uses of accomplishments in the progressive' (2000, 315-6).

3.5.4 The role of input: frequency and prototypes

While the aspect hypothesis studies cited above were concerned with learner production, researchers have also looked at the input learners receive. In one recent study for example, an analysis of 110,000 words of instructional input or 'teacher talk' in English language classrooms revealed that 72% of the verbs used in the past tense were telic, and 59% of the verbs used in the progressive were activity verbs (Collins et al., 2009) . This is in line with the distributional bias hypothesis, proposed by

Andersen (1990) and subsequently supported by a number of empirical studies. Native speakers in normal interaction with other native speakers, and with learners, tend to use each verb morpheme with a specific class of verbs in apparent conformance – or at least partial conformance – with the aspect hypothesis.

According to Andersen and Shirai (1994) this on its own could possibly account for the tense-aspect patterns found in interlanguage as learners exposed to native speaker input would initially interpret the skewed distribution of forms as an absolute characteristic of the forms themselves. However both Andersen and Shirai (Andersen, 2002; Shirai and Andersen, 1995) have posited that it is also possible that native speakers and learners are choosing particular verb-morpheme combinations for the same cognitive and discursive reasons, such as, for example, a predisposition to find prototypes:

> Humans naturally form prototypes of meaning and in learning new form-meaning relations will first assign the core prototypical meaning to a form (a verb, a verb construction or a gram), the meaning of which is inferred from the input based on context, through repeated noticings of the form. It then takes time for the novice learner to perceive other tokens of the verb or gram in the input that have less-prototypical interpretations. With time the novice learners gradually expand their mental representation of the meaning and distribution of the gram with different verbs. (Andersen, 2002, 93)

Noting that full targetlike use of past morphology requires that its use is extended to other environments as well, Bardovi-Harlig (2000) suggested that research into how interlanguage expands prototypes is 'particularly worthwhile' and, drawing on the work of Andersen, Shirai and others up to that time, she gave hypothesized paths for the progressive and the past with the sequences extending into the realm of pragmatics:

The spread of the past: Deictic past (achievement> accomplishment > activity > state > habitual or iterative past) > counterfactual or pragmatic softener.

The spread of the progressive: Process (activity > accomplishment) > iterative> habitual or futurate> stative progressive >pragmatic softener.

(2000, 429)

More recently, in her analysis of the progressive's occurrence in two L1 corpora, the BNC and the Bank of English, Romer (2005) grouped iterative and habitual together as 'repeated'. She found that in fact the progressive is used around 39% of the time to report actions or events which are repeated, as opposed to single events or actions. And she has suggested that 'repeatedness' should move further to the centre of descriptions of the English progressive. (Huang had also grouped iterative and habitual together in her 1999 study of oral data from three L1 speakers and five Chinese learners of English. She found as expected that both groups were more likely to use the progressive with durative verbs, but she also found that her L1 speakers tended to use the form to describe repeated events, while the learners used it mainly to describe the ongoing continuous nature of unitary situations).

Research which looks at the relationship between input and acquisition has the potential to shed light on specific areas of interest, like the aspect hypothesis, but also on the more general debate about the nature of linguistic knowledge. Ellis and Collins (2009) for example, argue that evidence of sensitivity to input frequency would entail that language users must have registered patterns of occurrence in processing, lending support to emergentist accounts of acquisition which emphasise the role of input.

Category Learning

In recent years there has been a growing body of research undertaken from functional, cognitive linguistic, psycholinguistic and constructionist perspectives, investigating how the learner's general cognitive processes interact with the input in ways which either promote or constrain the acquisition of particular constructions. Constructional acquisition can be viewed as, in essence, a process of categorisation (Goldberg, 2006) and a number of these studies have examined the degree to which acquisition can be explained in line with the principles of category (or concept) learning , key aspects of which include the frequency and frequency

distribution of exemplars, their salience and significance, and the match of their meanings to the construction prototype.

As Rosch et al. (1976) demonstrated in relation to the acquisition of nouns, people will categorise *robin* as *bird* faster than they will exemplars like *goose* or *penguin* which have less common features or feature combinations. Children acquire basic category terms like *bird* and *dog* before superordinates like *animal* or subordinates like *wren* and *poodle*. It has been suggested that learners acquire a new category or concept more easily if they are initially exposed to a low variance sample with obvious shared attributes, rather than a high variance sample which may obscure commonalities.

Boyd and Goldberg (2009, 421) say that 'low variance input has been shown to play a facilitatory role in the initial induction of constructional categories in language' and they comment that 'fortunately many aspects of the linguistic input appear to be naturally structured so as to reduce [...] variability'. One such aspect is the Zipfian distribution of words. Zipf's Law (1935) states, that in a corpus of natural language utterances, the frequency of any word is inversely proportional to its rank in the frequency table. The most frequent word, for example, will occur approximately twice as often as the second most frequent word and three times as often as the third most frequent word. As Boyd and Goldberg point out this means that although individuals may have upward of 60,000 words in their vocabulary, the number of words that they use with any regularity will be at least an order of magnitude smaller. 'This constitutes a significant decrease in input variability: Learners tend to hear the same set of high-frequency, high utility words over and over again' (2009, 421).

Researchers have looked at the roles of both low variance input and high type frequency in the learning of categories like 'past' 'progressive' and 'perfect' . Wulff et al. (2009), for example, investigated how various

features of the input – including frequency distributions and prototypicality of lexical aspect – affect the acquisition of temporality. As part of their first study they retrieved verb form frequencies for all verbs from two native speaker English corpora – the general BNC spoken corpus of 10 million words and the more specialised academic discourse corpus, MICASE, consisting of 1.7 million words. When they looked at the 100 most common verbs in each tense-aspect category they established, as expected, that the frequency distribution follows Zipf's Law. They found that a small number of common high-utility verbs dominate the 'top ten' list for each tense-aspect category, with six verbs – *be, do, have, say, get, go* – appearing in all four categories for the BNC corpus (Table 3.1). Given this overlap they concluded that it is unlikely that the acquisition of tense-aspect can be driven solely by raw frequency of occurrence.

Wulff et al. considered the possibility that learners may also be sensitive to how 'distinctively' a verb is associated with a particular tense-aspect. To establish the most distinctive verbs in each category Wulff et al. did a multiple distinctive collexeme analysis (MDCA) which calculates a verb's degree of bias toward (or away from) a given construction as compared to a fixed set of alternative constructions. This produced 'top ten' lists which were quite different from the raw frequency ones – in fact only two verbs are found on both BNC lists : *call* for perfect, and *come* for progressive (Table 3.2). As Wulff et al. point out, the rankings for the MDCA generated lists 'reflect intuitions about verbs that typically occur with different TA [tense-aspect] categories: The past and perfect TA columns are occupied by highly telic verbs such as *die*, *crash*, *explode*, *lose*, or *finish*; the progressive (unsurprisingly) prefers continuous action verbs like *sit*, *play*, *walk*, and *run*' (2009, 362).

TABLE 3.1

The 10 Most Frequent Verbs by Tense–Aspect Category in BNC_{spoken}

Past	Perfect	Present	Progressive
be	**get**	**be**	**go**
do	**be**	**have**	**do**
have	**do**	**do**	**be**
say	**have**	know	**say**
get	**go**	**think**	**get**
go	**see**	mean	**come**
think	make	**say**	**have**
come	call	**go**	look
take	**take**	**get**	talk
want	**say**	**see**	try

TABLE 3.2

The 10 Most Distinctive Verbs by Tense–Aspect Category in BNC_{spoken}

Past	Perfect	Present	Progressive
start	lose	suppose	look
die	leave	need	come
become	finish	want	sit
wake	square	bet	play
crash	rid	seem	wait
retire	allow	excuse	walk
panic	base	hope	joke
explode	bear	tend	run
disappear	call	reckon	watch
occur	marry	sound	deal

from Wulff et al. 2009

When they looked at the 100 most distinctive verbs in each tense-aspect category, Wulff et al. found that the association-based ranking of these verbs is also Zipfian and they suggest it follows that these verbs could provide the category-specific input that learners need to acquire the semantic restrictions of the different tense-aspect categories.

Wulff et al. also tested the hypothesis that the verbs most distinctively associated with each tense-aspect construction in the input are prototypical of the meaning of that construction. In order to investigate the prototypicality of the verbs, native speaker telicity ratings were obtained for a range of verbs selected from these analyses. 20 native speakers of American English were given a questionnaire which presented the verbs in isolation, without arguments, and in their base forms; they were asked to evaluate each verb in terms of how strongly it implied an endpoint. The resulting data demonstrated that those verbs distinctively associated with past tense in the input received significantly higher telicity ratings than verbs associated with the progressive.

The final hypothesis tested by Wulff et al. was that the first-learned verbs in each tense-aspect construction are prototypical of that construction's functional interpretation in terms of their telicity/lexical aspect. For this

part of the study, oral production data from Bardovi-Harlig (2000) was analysed. The data in question was collected from 37 beginning L2 English learners, from five different L1 backgrounds, who were asked to retell a film excerpt in their own words. The average number of verb tokens per learner narrative was 51. From this data set the researchers selected verbs that occurred more than 10 times overall and which were distinctively associated with present, past tense , or progressive as determined by a chi-square test. The five most frequently occurring past tense verbs in the learner production data, and the five most frequently occurring progressive verbs differed significantly in their mean telicity ratings, with the past tense verbs being judged more telic, and the progressive verbs more atelic.

According to Ellis (2013a) the results reported in Wulff et al. suggest that the verbs first learned by adults in the progressive are also frequent in the progressive in the input, distinctively associated with the progressive in the input, and highly atelic. Likewise, the verbs first learned in past tense are frequent in past tense in the input, highly distinctive for past tense in the input and highly telic. Ellis commented that the study's findings provide some support for the hypothesis that the learning of tense and aspect, like that of other linguistic constructions, can be understood according to psychological principles of category learning: 'TA construction is sensitive to input frequency, reliabilities of form-function mapping and prototypicality of lexical aspect in English' (2013a, 104). Ellis also noted, however, that although the study involved quite extensive corpus analysis 'it is a stretch to claim that the language sampled therein was properly representative of that to which the ESL learners had been exposed. Additionally the learner data was small, far from dense, and it covered only a very short period of initial acquisition' (2013a, 104).

The ways in which frequency, frequency distribution and prototypicality of meaning affect acquisition has also been examined in relation to verb-argument constructions (VACs). Using data drawn from the seven

English learners in the ESF study, Ellis and Ferreira Junior (2009) examined three VACs – verb locative, verb object locative and ditransitive. They found that the most frequent, most prototypical and most generic verbs for each construction were those learned first (*go somewhere* for verb locatives; *put something somewhere* for verb object locatives; *give someone something* for ditransitives) and that the frequency profiles of these verbs were Zipfian. Interest in VACs has grown in recent years and Ellis et al. (2013) report on a major research programme which plans to analyse hundreds of verb constructions in native speaker English using the BNC corpus of 100 million words. Preliminary findings from studies of 23 VACs suggest, they say, that zipfian distributions make language 'robustly' learnable from input. When they ranked verbs in each construction by frequency they found a zipfian distribution, but perhaps more significantly, the rankings of verbs in particular VACs differed 'markedly' from the rankings of verbs in the language as a whole: ' VACs are selective in their verb constituency' (2009, 36). The preliminary study also measured contingency, following Shanks (1995) observation that the more reliable the contingency between a cue and an outcome, the more readily an association between them can be learned. Using 'directional mutual information' and 'faithfulness' measures, they found a general pattern which would seem to aid acquisition, with individual verbs tending to select particular constructions and particular constructions selecting particular verbs.

In addition to corpus analysis, a number of studies have also analysed the processing of VACs by native English speakers and L2 learners. Ellis et al. (2014) used free association tasks to investigate L2 verb-argument constructions and the ways in which their access is sensitive to statistical patterns of usage. 131 German, 131 Spanish, and 131 Czech advanced L2 learners of English were asked to give the first word that came to mind to fill the V slot in 40 sparse VAC frames such as 'he ___ across the....', 'it ___ of the....', etc. For each VAC, the results were compared with corpus analyses of verb selection preferences in the BNC 100 million word

corpus and with the semantic network structure of the verbs in these VACs. Ellis et al. report that for all three L1 groups, multiple regression analyses predicting the frequencies of verb types generated for each VAC show independent contributions of (i) verb frequency, (ii) VAC-verb contingency, and (iii) semantic prototypicality. In other words, which verbs come to mind when the advanced L2 learners consider the prompts are determined by verb token frequencies in those VACs in usage experience; by how faithful verbs are to particular VACs in usage experience, and by the centrality of the verb meaning in the semantic network of the VAC in usage experience.

Romer et al. (2014) used the same L2 data set as Ellis et al. (2014). In order to determine how similar or different learners' verb–VAC associations are from those of native speakers, they compared evidence of the learners' mental representations of 19 VACs with data collected from 131 L1 English speakers performing the same task. They reported that while there was some overlap between learners' and native speakers' mental representations of VACs, there were also differences in the associations of verbs and constructions. For example, they found that all three learner groups relied more on general, highly frequent verbs (e.g. *be, come, do*) and produced lower numbers of specific, less frequent verbs (e.g. *slip, reach, crawl*) than native speakers did. They also found that the German L1 and Czech L1 learners produced more verbs that correlated more closely with those produced by the L1 English speakers than the Spanish L1 learners did. Romer et al. suggested that this difference may be due to the fact that languages differ in the ways in which verb phrases express motion event and in this respect, German and Czech are typologically closer to English (all three being satellite-framed languages) than Spanish which is a verb-framed language.

In relation to these studies of L2 processing of English VACs, Ellis et al. (2014) observed that while their findings reflect L2 knowledge of language that comes from usage, second language constructions reflect usage of

both L2 and L1. Because L2 learners have devoted considerable resources to the estimation of the characteristics of the L1 in which they have become fluent, their computations and inductions are often affected by transfer, with L1-tuned expectations and selective attention blinding the acquisition system to aspects of L2 form and meaning.

A number of other studies have looked at the role played by input frequency effects in second language acquisition. Collins et al. (2009), for example, addressed the question of what makes some grammar 'easy' to learn and other grammar 'difficult'. They performed a multidimensional analysis on a number of constructions as they occur in instructional input; the input in question was a 110,000 word corpus of 'teacher talk' to young learners of English, and the constructions included the past –ed and the progressive –ing. They found that 70% of progressive marked verbs came from the 'top 1000' word families, and that 26 of these verb types occurred with *–ing* at least 8 times (and some much more than that, with *doing,* for example, being used 69 times and *talking* 54 times). In comparison, only 58% of verbs marked for regular past came from the 'top 1000' and, more significantly, only 4 of these occurred more than eight times, with the most frequent, *asked*, being used 17 times. Collins et al. suggest this shows that opportunities to learn exemplars of the progressive far outweigh opportunities to learn regular Simple Past. They say that this could explain the earlier acquisition of the progressive, '[a]ssuming that eight instances of listening exposure to a particular verb and its tense marker result in a lasting mental association' (2009, 344) (they chose eight as a baseline following the findings of studies of incidental vocabulary acquisition in reading, although they acknowledge that acquisition from spoken input would probably require exposure to a larger number of instances).

Boyd and Goldberg (2009) comment that low-variance skewed input and high type frequencies can both play a role in the acquisition of constructions – the former in relation to the initial creation of categories

and the latter in relation to generalisation, to the development of the abstract representations needed to support productivity. They say that these two properties of input are not necessarily opposed to each other - even though that may at first seem to be the case - since increased type frequency generally goes with increased variability. However a Zipfian distribution of verbs within a construction allows for high type frequency and low category variability to exist side by side in the same sample, and they say that their own research indicates that while Zipfian input draws attention to the handful of high frequency verbs, learners are still exposed to sufficient different verb types to drive abstraction and generalisation processes.

3.6 The acquistion of future temporality

3.6.1 Introduction

According to Ayoun (2014, 181), the investigation of future temporal reference is a neglected area of L2 acquisition research, with the vast majority of empirical studies focusing on past temporal reference. This is certainly the case with English where there is a substantial body of work dealing with different aspects of the L2 acquisition of past expression, but relatively little attention paid to the way L2 learners acquire the ability to talk about the future – to make predictions, discuss plans and express intentions in English. The principal studies in this area were carried out by Bardovi-Harlig (2002, 2004a, 2004b, 2005) and her research on the use of lexical futures and on the two main grammaticalized forms, the Will-Future and the Go-Future is reviewed in section 3.6.3. This section begins however with a brief overview of what typological studies and corpus based research have established about L1 patterns of use in this area of the tense-aspect-mood system.

3.6.2 Future time reference in English

Because the future is not fact – because, as Dahl (2000a) puts it, we cannot perceive or remember future states of affairs - future time talk encompasses not only temporality but also modality. It may concern intentions relating to the future, in which case, Dahl says, it is by definition restricted to things which are under our control (or are believed to be) and prototypically occurs in sentences with a human subject. We may also make predictions about the future, most typically concerning courses of events that are not within human control or at least not within the control of the speaker.

According to Bybee (2010) the development of future tense markers usually involves a stage in which the markers express intention. This meaning of intention cannot come about by semantic generalization, she says, but rather must be attributed to frequently made inferences, a hypothesis which she says is supported by cross-linguistic reports of an intention use for futures that evolve from various sources – movement towards a goal, volition, obligation and even temporal adverbs.

In English, *will* originally meant 'want' but its meaning was extended over time from volition and intention to prediction and to its functioning as a general marker of the future. *be going to* started out as an expression which simply indicated that someone was literally going somewhere to do something, and then over time, through a metaphorical extension, took on the meaning of intention to do. Of these two grammaticalized future forms, *will* developed first, and continues to dominate English future expression, providing the language with its nearest approximation to a neutral or colourless future (Leech 2004). *be going to* arrived relatively late – according to Beckner et al. (2009) it was not until the 17th century that it spread among English speakers.

The Will-Future is used in a wide range of contexts in which it is appropriate to make predictions, Leech notes, and is particularly common in the main clause of conditional sentences. The form can refer to a definite or indefinite time in the future and is suitable for both 'long-range and short-range forecasts' (2004, 57).

According to Dahl (2000a) the Western European de-andative *(go)* construction tends to be used for intention-based non-remote future time reference, but in some languages it has extended its reach to non-intentional contexts, especially imminent ones. This is the case with *be going to*. Leech identifies 'the future outcome of present circumstances' as the general meaning which L1 speakers attach to the form. Apart from expressing present intentions (*I'm going to lose weight and get fit*) the Go-Future can also express predictions based on present cause (Look at those clouds *–it's going to rain)*. And while the form is often used in reference to the immediate future, it can be used to refer to a more remote one. Dahl observes that it is common for there to be more than one future form moving along the same grammaticalization path and that in such situations, the forms will have partly overlapping functions and will compete with each other. Leech reports that in more informal styles of English (particularly in speech) *be going to* is beginning to rival *will* as a fairly neutral future auxiliary.

Like the Go-Future, the Present Progressive is used in English to refer to a future happening anticipated in the present, but in this case it is an existing plan or arrangement that is signalled (*we're going to the theatre tonight*). Leech (2004, 62) notes that the difference between 'arrangement' and 'intention' is a slight one, and that *be going to* can be substituted for the Futurate Present Progressive, but that when it is, there is a subtle change of emphasis 'An intention is part of one's present state of mind, while an arrangement is something socially predetermined in the past, regardless of how the speaker feels now'. Romer (2005) found that in the BNC and Bank of English corpora the form is closely

connected with dynamic action, supporting Leech's observation that it is mainly found with doing verbs involving conscious human agency. Dahl (2000a) points out that the examples cited in the literature tend to involve movement verbs, but he suggests that it is not the sense of 'movement' but the sense of 'preparation' which is relevant to the form. This observation is supported by Nesselhauf and Romer (2007). In an analysis of the *spoken* BNC corpus they found that the form is possible with a large number of verbs, rather than being restricted or largely restricted (either in absolute or in relative terms) to one semantic group, such as motion verbs. However they also found that *go* and *come* are by far the most frequent verbs found in the Futurate Present Progressive. Dahl notes that there may be a tendency in English to avoid using *be going to* with movement verbs and that this favours the use of the progressive with those verbs. In fact the practice is common enough to feature in pedagogical grammars and, as Parrott (2010) points out, this means that the Futurate Present Progressive when used with *go* or *come*, for example, may not always be expressing an arrangement, but sometimes the ' intention' meaning associated with the Go-Future.

A number of other forms can also occur in English with a future meaning. The Simple Present is used in dependent clauses when the future time reference is clearly established in the main clause (*If I see John I'll tell him*). This Subordinate Future Simple Present is found in dependent clauses introduced by conditional or temporal conjunctions, and it can also follow that-clauses where it is typically used in the imperative (*Make sure you get enough sleep tonight*). A much less frequent use of the Simple Present is in independent clauses where it expresses 'a future assumed to be fact' (*Tomorrow's Saturday*) as well as ' a plan or arrangement regarded as unalterable' (*We start for Istanbul tonight*) (Leech, 2004, 65).

will +progressive infinitive is typically used to refer to temporary or ongoing situations in the near, but not too immediate future. It also has a

special use which Leech calls the 'future as a matter of course'. Here the form appears to combine the 'prediction' meaning of *will* with the 'arrangement' meaning of the Futurate Present Progressive. Leech speculates that this use may have developed as a means of avoiding the overtones of volition which *will* sometimes combines with its prediction meaning. Because it suggests that something will be happening anyway, as a result of an existing arrangement, *will* +progressive infinitive is often a more polite and tactful alternative to the Will-Future, he says, and this is probably why it is becoming more common in everyday speech.

Leech (2004, 69) gives the following order of frequency for the main future forms in English:

1 will +infinitive
2 Subordinate Future Simple Present
3 be going to +infinitive
4 Futurate Present Progressive
Futurate Simple Present
5 will + Progressive infinitive

Infrequently occurring forms include the *be about to* and *be to* constructions. *be about to* refers to the immediate future and is close to the meaning of *be going to,* except that it suggests greater immediacy: *Fasten your seat belts. We are about to land.* The *be to* construction is usually used in formal written English to refer to something which will happen as a result of a decree or plan, normally made by some one, or some authority other than the speaker of the sentence.

Finally, Leech notes that all these constructions are variations on the present tense, using present tense forms or present tense auxiliaries. The present tense in a broad sense encompasses both present and future he says, and this helps to explain why English, lacking a future tense as such, uses present forms to express future meanings.

3.6.3 SLA research and future time expression

As already noted, while many researchers have looked, from a variety of perspectives, at the L2 acquisition of past tense-aspect in English, there have been relatively few studies of the development of future time expression. Bardovi-Harlig has published four - two on the Will-Future and Go-Future (2004a, 2004b) one on lexical futures (2005) and one on formulaic use (2002). Her analyses are based on the same samples used in her studies of past-related morphology (1997, 2000), in other words on 1576 written and 175 oral texts collected over a two year period from 16 college students enrolled in an intensive English language programme (the mean length of observation was 11.5 months, ranging from 7 to 17.5 months).

Bardovi-Harlig (2004b) reports that her learners' future expression is dominated by *will*, which accounts for 55% of the 2,566 future contexts in the written data, and 63% of the 1,169 future contexts in the oral data. Lexical futures *(want to, have to, hope)* come second at 20% and 15% respectively, followed by *be going to* at 9% and 7%. Finally there are present tense forms used with future meaning - the futurate present progressive, for example, accounts for 2% of the written, and 1% of the oral contexts.

Bardovi-Harlig's examination of the emergence and developing use of verbs such as *want to, have to* and *plan to* (which she refers to as 'lexical futures') was motivated in part by evidence from longitudinal studies demonstrating that learners pass through various stages in the acquisition of temporal expression, with a lexical stage preceding the final, morphological stage of development (Dietrich et al., 1995). Bardovi-Harlig notes that:

> On the one hand, this cycle of lexical-before-morphological stages might be expected to extend to lexical futures because lexical futures encode the future semantically rather than

> morphologically. On the other hand, lexical futures are different from the lexical devices included in interlanguage analyses to date, namely, adverbials, locatives, and nominals which co-occur with the verb, but not the verb itself. (2005, 2-3)

She reports that the lexical future and *will* emerge 'similarly early' in the corpus, '[h]owever, emergence—the very earliest stage—does not sufficiently distinguish between *will* and the lexical future. The longitudinal view must be taken into account.' (2005, 8). There appeared to be multiple patterns in the interlanguage development of these learners, she says, with two learners using *will* as essentially their only marker of future, five using both *will* and lexical futures, and the remaining nine using *be going to* in addition to *will* and lexical futures. The data for five of the learners showed an eventual reduction in lexical futures with the establishment of *be going to*. Noting that the use of lexical futures persisted for most learners throughout the observation period, Bardovi-Harlig comments that this was in keeping with target language norms, 'since lexical futures are a part of future expression cross-linguistically' (2005, 11). Desire was the most common meaning expressed by lexical futures in Bardovi-Harlig's data, with *want to* accounting for 72% of all tokens (533/733), followed by obligation - *have to* accounting for 14% (102/733).

Focussing on the two major grammatical forms, Bardovi-Harlig (2004b) reports that *will* outnumbered *be going to* 5.8:1 in the written sample and 9.3:1 in the oral sample. In comparison, she says, Longman's Corpus Grammar reports a ratio of 2.5:1 in spoken British English, and 1.6:1 in spoken American English. 'This comparison suggests the degree to which *going to* is under-represented in the interlanguage of these learners' (2004b, 182).

Bardovi-Harlig established an order of emergence, with *will* emerging early – the majority of the learners show emergent use before the middle of the second month, and all before the end of the third month. For 14 of the 16 learners *will* clearly emerged before *be going to*; for the remaining

two learners both forms emerged in the same 15 day period. Seven of her learners produced less than 10 tokens of *be going to*, and 9 produced 22 or more tokens. Two of what Bardovi-Harlig calls the 'high' producers mainly used the form for introductions to compositions or oral presentations. *I am going to write (about)* was used by Satoru 29 times in the written data and by Khaled 24 times. Four other learners also used the form in this way and Bardovi-Harlig says that for five of these six learners most of this use seems to be formulaic, but the formula does not account for their first use of *be going to*. Commenting on whether this use of the form for introductions 'opens the door' to further development, she notes that while Satoru for example showed no development of productive use in his written data, Khaled did.

Bardovi-Harlig reports that in her data the main meaning associated with *be going to* in the early stages is plans that have been made or are underway, that *be going to* and *will* may then occur in complementary distribution and that the later stages seem to be characterized by a broader use of *be going to*. She gives the example of Guillermo whose 25 written tokens included 8 introductions (7 of them *I am going to write*) 16 plans and schedules, and one prediction. It is, she says 'a characteristic of later stages, or more advanced learners, that they generalize the use of *going to* to predictions' (2004b, 192). Some learners however did not settle on a particular meaning for *be going to*. She gives the example of Zayed who used the form 6 times in his written samples, 3 times for prediction, twice for introductions and once to express a resolution.

A study of the association of grammatical person with particular future forms is 'somewhat thwarted' in her corpus, Bardovi-Harlig says because the learners often wrote and talked about themselves, and therefore used first person. In her opinion, however, 'because of the high degree of intentionality and volitionality associated with *going to*, it is not surprising that the initial subjects are first person' (2004b, 193). She says

that when the learners start to use the form to discuss the scheduling of events this allows an opening for the 3rd person, as does the use of the form with predictions, when they appear. *will* has a higher proportion of 3rd person subjects she adds, but does not elaborate on this.

Bardovi-Harlig also notes that while her corpus lends itself to an examination of intention- and prediction-based future time reference, it is 'less well-suited' to the study of remoteness because learner journals and conversational interviews generally considered the near future and only rarely touched on the remote future (2004b , 198).

Bardovi-Harlig suggests that it is likely that multiple factors contribute to both the early emergence of the Will-Future and its frequency of occurrence relative to the Go-Future (2004a). In her discussion she considers formal complexity, as *be going to* involves more phonological units and is inflected for person, number and tense. However she points out that all her learners clearly enter the morphological state of development and are already using other compositional tense-aspect forms. For all learners the Past Progressive emerges before the *go*-future, and the Present Perfect emerges around the same time. So, she concludes, while formal complexity may explain why *will* emerges first it does not explain why *be going to* is comparatively delayed or infrequent in the production of half of the learners in her study.

Given that learners have been shown to use lexical means - like adverbs - to express temporal relations before they acquire the morphological means to do so (Dietrich et al., 1995), Bardovi Harlig suggests that *will* might have an advantage over *be going to* if learners regard it as a lexical marker of the future rather than a grammatical one. In support of this, she points out that *will* is a free morpheme and that in her learner corpus it ' virtually lacks variants', there being very few tokens of the phonologically reduced ('ll) form (2004a, 140). Furthermore *will* emerges early and 'the timing of the emergence of *will* patterns more

similarly to lexical time devices than to morphological ones' (2004b , 184). Finally she considers the role of input, noting that *will* is more than twice as frequent as *be going to* in oral conversation and that *be going to* is 'reasonably rare' in written input:

> While these factors conspire to assure that learners associate *will* with the future, the one-to-one principle (Andersen 1984) - under which learners associate one form with one meaning – assures that the result is a fairly stable system of future time reference wherein *will* is the dominant form. (2004b, 184)

Bardovi-Harlig says that the strength of the one-to-one principle can be gauged by the fact that half of the learners in her study did not show productive use of *be going to*. In those cases where the form emerged and became established, learners began with a limited form-meaning association, mainly using the Go-Future to express intentionality, specifically plans that have been made or are underway – a meaning initially (and, to some extent, subsequently) encoded by the Will-Future. She suggests that in this way '*going to* takes on a decidedly modal flavour to break into future reference before generalizing to include predictions' (2004b, 196-7).

Part Two RESEARCH DESIGN & METHODOLOGY

4 Research Design & The Research Context

4.1 Introduction

This chapter begins by briefly outlining the theoretical framework within which the study is conducted (a more detailed discussion of Emergentist perspectives is provided in 2.2 and 2.3). The research questions are presented in 4.3 and some issues related to the research design are then discussed in 4.4: the rationale for conducting a longitudinal study (4.4.2) the importance of sampling implicit - as opposed to explicit - linguistic knowledge (4.4.3) and the choice of written rather than oral production data (4.4.4).

The research and learning context is discussed in 4.5. As Ellis and Barkhuizen (2005) among others, have pointed out, in a study of the kind undertaken here it is important that the researcher provide detailed information on the learners and their learning context and - in relation to data collection - the conditions under which the learners are performing, and the kind(s) of language use sampled by that performance. Information is provided in 4.5.2 on the school environment and curriculum, and in 4.5.3 on the language support programme to which the students belonged. 4.5.4 gives a detailed description of the writing process from which the data was drawn, outlining the ways in which the process was designed to encourage relatively spontaneous production. Profiles of the three learners are provided in 4.6, including information about the learners' language backgrounds and their patterns of exposure to English outside the classroom.

4.2 Theoretical framework

This study is conducted within an Emergentist framework, the central thesis of which is that 'simple learning mechanisms, operating in and across the human systems for perception, motor-action and cognition as they are exposed to language data as part of a communicatively-rich human social environment by an organism eager to exploit the functionality of language, suffice to drive the emergence of complex language representations' (Ellis, 1998, 657). A usage-based theory of grammar is adopted in which the cognitive organization of language is based directly on experience with language, the basic units of language being constructions : form-meaning pairs of various levels of abstraction, and various sizes, from morphemes and single words to whole clauses, each of which serves some communicative or socio-pragmatic function (Croft, 2001; Goldberg, 2006; Tomasello, 2003). In L2 acquisition (as in L1 acquisition) input and interaction are regarded as central to construction learning (Beckner et al., 2009) with frequency being a key determinant of acquisition because 'language knowledge involves statistical knowledge, so humans learn more easily and process more fluently high frequency forms and 'regular' patterns which are exemplified by many types and which have few competitors' (Ellis, 2013a, 89). As a result construction learning is affected by a range of factors: factors relating to the form, such as frequency and salience; factors relating to the interpretation, such as significance in the comprehension of the overall utterance, prototypicality, generality, and redundancy; factors relating to the contingency of form and function; and factors relating to learner attention, such as automaticity, transfer and blocking (Boyd and Goldberg, 2009; Ellis and Collins, 2009; Ellis and O'Donnell, 2012; Goldberg, 2006; Luk and Shirai, 2009; Romer et al., 2014; Tomasello, 2003) .

The present study investigates how form-meaning associations are established and then revised in the course of L2 language development. In the discussion of the findings in relation to the research questions, the possible influence of frequency, salience, prototypicality and contingency factors will be considered. While an analysis of cross linguistic influence is beyond the scope of this study, some reference will also be made to L1 transfer and blocking.

4.3 Research questions

This study examines aspects of the developing English language of three teenage L2 learners and investigates whether, and to what extent, development continues up to the end of the learners' secondary school education in Ireland. Temporality was chosen as the main focus of the study because of the important role it plays in linguistic communication.

The following research questions were selected :

Past-related temporal expression: Does the Present Perfect emerge in the interlanguage of all three learners? Does the Present Perfect emerge after the Simple Past and Past Progressive, as reported by Bardovi-Harlig (2000)? If so, does the emergence of the Present Perfect depend on the stability of the Simple Past, as reported by Bardovi-Harlig (2000)?

In the case of the Simple Past, Past Progressive and Present Perfect, is there evidence of continuing development in the use of these forms – of their spread from typical to less typical meaning associations and uses?

Future temporal expression: What future forms emerge in the learners' interlanguages, and what is the order of emergence of these forms? Does the Will-Future (*will* +infinitive) initially dominate future

time reference, with the Go-Future *(be going to)* emerging late, as reported by Bardovi-Harlig (2004b)?
If the Go-Future emerges late, is there evidence in the data to suggest which factor(s) might be responsible for its late emergence?
What meanings are initially associated with emerging future forms? How are form-meaning associations revised following the emergence of new forms?

For this study, supplementary information on the learners' general language development was also obtained in relation to lexical diversity and overall accuracy of use. The research questions were:
Does overall accuracy of language use continue to increase?
Does lexical diversity continue to develop and does it approach L1 norms during the course of the study?

4.4 Research design

4.4.1 Introduction

Since the 1970s the study of second language acquisition has evolved and expanded to embrace a wide range of research disciplines and interests, employing a variety of approaches and engendering much debate about the relative merits of different research paradigms and methodologies. As Chapter 3 illustrates, the study of tense-aspect morphology, for example, has involved a diverse range of disciplines from computer modelling to corpus analysis, and has generated dozens of different studies - some longitudinal but most cross-sectional; some experimental hypothesis-testing inquiries , others non-experimental descriptive studies, some presenting qualitative analyses of their results, others using quantified analyses. It would appear then that anyone embarking on research in the field of temporality has a range of options to consider in terms of research design and methodology, but, as pointed

out earlier (2.4), every choice involves compromises... not least in terms of validity.

For Larsen Freeman and Long (1991, 14) : 'what is important for researchers is not the choice of *a priori* paradigms or even methodologies, but rather to be clear on what the purpose of the study is and to match that purpose with the attributes most likely to accomplish it. Put another way, the methodological design should be determined by the research question'. Dornyei (2007) generally agrees that the best method is the one which answers the research questions most efficiently, but he poses the question 'how do we know in advance what will work?' In most cases he says, a number of approaches might be suitable for exploring a problem, but sometimes there may be an obvious match between a research question and one method of inquiry. He points out, for example, that well established research traditions exist in some areas, and if it is possible to identify such a 'methodological salience' associated with a particular question or issue then it can make sense to follow in that tradition – not least because results may then be compared with others' findings.

Working from an emergentist perspective, this researcher has undertaken to investigate aspects of the acquisition of tense-aspect morphology by some learners of English. This involves tracking the emergence and developing use of a number of past and future forms in relation to the development of other forms in the temporality system, identifying patterns of change and looking for explanations of cause and effect. An examination of the methodologies employed in previous temporality studies and in particular the studies most relevant to this one (e.g. Bardovi-Harlig 2000, 2004b) suggested that the most appropriate design to employ here was a descriptive non-experimental quantitative longitudinal study.

Ortega and Iberri-Shea provide a useful summary of the main features of this approach:

> Such studies focus on quantifiable variables, but without researcher's manipulation and without random sampling or random assignment of participants to conditions or to contexts for learning. They are characterized by multiwave data collection from the same individuals over a relatively long period of time, which typically spans anywhere between four months and four years. Although the longitudinal data in these studies are quantified, the number of participants is far too small to warrant the use of inferential statistics. For this reason, descriptive statistics displayed in the form of frequencies, percentages, and proportions, and other analytical tools, such as visual displays and implicational scaling, are favoured. (2005,29)

4.4.2 A longitudinal perspective

In their review of trends in second language research, Ortega and Iberri-Shea (2005, 26) comment that: 'Both common sense and expert knowledge tell us that learning a language other than the mother tongue [..] is a complex process that happens through and over time.' It can be argued, they say, ' that any claims about 'learning' (or development, progress, improvement, change, gains, and so on) can be most meaningfully interpreted only within a full longitudinal perspective'. Other reviews of SLA research have also asserted that the effective study of language *development* requires longitudinal methods (e.g. Mellow et al., 1996). DeKeyser (2014, 367) notes that longitudinal studies are 'fairly rare' in the field, even though 'longitudinal research is ideal for answering questions that address long-term development, as most of the questions in SLA do' .

Cross sectional studies predominate in SLA research and they have an important contribution to make, especially in relation to hypothesis testing. However because a cross sectional approach quintessentially involves 'a snapshot-like analysis of the target phenomenon at one particular point in time' (Dornyei 2007, 78) it is arguably ill-equipped to capture the developments which are the main subject of this particular

study - the emergence of a tense-aspect form in the learner's interlanguage and the various stages traversed on the way towards its acquisition. This is especially so because the learner's use of a particular form is affected by the emergence and development of other forms, with form-meaning associations being revised again and again until the entire tense aspect system has been acquired (see Section 3.5 for a discussion). As a result, rates of appropriate and accurate use fluctuate, and while a cross sectional 'snapshot' might capture the apparently target-like use of a form at one particular point in time, it may tell us little about that form's actual stage of development in the learner's interlanguage. Bardovi-Harlig (2000) points out that it is possible to design a cross sectional study in such a way that it does provide useful information on acquisition sequences – by including a range of proficiency levels for example, and by coding form and meaning separately. However she herself chose a longitudinal approach for the studies on past morphology and future expression (2000, 2004a, 2004b) to which several of the research questions in this study refer.

According to Ortega and Iberri-Shea (2005) there is no general agreement on the optimal length of observation for a longitudinal study of bilingual development and they note that while an observation period of about one year was enough for the development of the English Past Perfect to be mapped in the case of 10 of the 16 learners in Bardovi-Harlig's 1994 study, another study (Long, 2003) reported that a 16 year observation span may not be sufficient for research on fossilisation. In their review of recent SLA research Ortega and Iberri-Shea found that studies ranged from as short as three months to as long as six years. Most of the studies involved learners who were either post primary or third level students and therefore an important factor in deciding observation span was probably 'convenient scaling' on institutional time – term length and course duration.

In the case of this study a four year time span was chosen for two principal reasons: (a) to establish whether or not English language development continued throughout the entire period of the learners' participation in the language support programme, and (b) to capture the development of late-emerging past and future tense-aspect forms.

In relation to the optimal timing (frequency and spacing) of measurements in a longitudinal study, Ortega and Iberri Shea suggest that 'this is in part a matter of grain size of the phenomenon under investigation' and that ' It would certainly be impossible to attain the level of detail needed to chart many L2 developmental phenomena without striving for frequent data' (2005, 39) . In the ground-breaking ESF studies of temporality the learners were followed for up to 30 months and the average distance between two recordings was about one month. Klein comments 'this may seem a long time [...] but the developmental process is so slow that the risk of missing important changes is almost non existent' (1995a, 31). In the case of L2 English learner Lavinia, from whom 14 samples were collected, the recordings were sometimes two months apart. In Bardovi Harlig's temporality studies, where the learners were enrolled in an intensive English language programme, three texts per learner were sampled, on average, in each two week period of term time over a period of up to one and a half years.

For this study 3 or 4 texts, on average, were sampled each school month (around 7 per sampling period) – the exact amount depending in each case on the number of spontaneously written texts a learner had chosen to produce in that month/period.

4.4.3 Sampling natural language use & implicit knowledge

As we cannot look directly into the mind of the learner, their underlying linguistic competence can only be examined by some kind of performance and a key methodological issue is what kind of performance

provides the most valid and reliable information about that competence. According to N.C. Ellis (2008a , 10) 'much existing research has taken a pragmatic approach and [..] used easy to administer grammaticality judgments, or metalinguistic judgments, or multiple choice or other limited response format measures of language proficiency'. N.C. Ellis is just one of a number of researchers who have questioned the validity of such measures. Rod Ellis and Gary Barkhuizen (2005, 21) for example, argue that the construct validity of a data collection method is best established by demonstrating that the performance it taps reflects, as far as possible, the kind of use for which language has evolved and is acquired. 'People do not usually go around matching pairs of sentences, matching sentences to pictures, or judging whether isolated sentences are grammatical or ungrammatical as part of their day to day lives'. Where researchers believe it is important to analyse 'naturally occurring' language use they can choose a non-experimental approach, collecting samples of the learners' speech or writing. However this is no straightforward matter because ' the language that learners produce can vary enormously depending on the type of production that is elicited'.

Accounts of L2 learning may differ in their theoretical perspectives but they generally agree that linguistic competence consists primarily of implicit L2 knowledge (Ellis, R., 2009) and it follows that a longitudinal study of acquisition processes should try to track the learner's developing implicit knowledge. But explicit knowledge may also be deployed in oral and written L2 production and this needs to be taken into account when deciding on elicitation measures.

Language learners' experience of the distinction between explicit and implicit knowledge was succinctly expressed by Ana, one of the participants in this study, when she said ' I do the grammar exercises, I get them right – so why I make mistakes when I write?' When she is focused on a cloze filling exercise or grammaticality judgement, or just taking particular care to speak or write accurately, when she has 'time to

think', the learner can draw on her explicit knowledge – on knowledge which is learnable as 'rules', is consciously held and available for controlled processing, and she can use this knowledge to remind her to stick *–ed* onto the verb for the Simple Past, or add *'s'* to the verb for the 3rd person singular Simple Present. However when the learner is focused on the immediate communication of some meaning she is more dependent on her unconsciously held implicit knowledge. As a result, if she has not yet acquired the Simple Past she may fail to inflect the verb, and if she has not acquired agreement she may leave out the 's'.

Where elicitation measures encourage or facilitate the use of explicit knowledge, the resulting production samples may suggest that the learner's acquisition of the morphology of the language is more advanced than is actually the case – given that we are chiefly concerned with a competence which is deployed in real-life, real-time situations.

Research suggests that the greater the focus on communication, the less likely the learner is to employ explicit knowledge (Ellis., R, 2008). According to Bardovi Harlig (2013, 220) open-ended tasks provide the best opportunities for observing interlanguage at work in the act of communication: 'The use of world knowledge, text structure, lexicon, lexical temporal expressions and verbal morphology come together in rich and varied discourse that reflects the voices of individual learners and their current level of development.' Bardovi Harlig also suggests that open-ended tasks are particularly suited to longitudinal studies: 'Because learners construct their own text, such tasks avoid the problem of learners learning the test as they might with various cloze passages or other item-focused tasks which have target responses.'

With the aim of sampling implicit rather than explicit knowledge, to as great an extent as possible, the written data used in this study was elicited through a process designed to encourage relatively spontaneous language use, a process involving open-ended tasks with a focus on

communication rather than a focus on the achievement of accuracy. (A detailed description of this process is provided in Section 4.5.4)

4.4.4 Reasons for choosing written data

While it is arguably easier to obtain samples of spontaneous language from oral recordings, the use of written data was considered preferable here for a number of reasons. As discussed in greater detail in Section 4.5.3 , the research was undertaken within the context of a post-primary school language support programme and one of the principal responsibilities of the programme was to help the learners access the mainstream curriculum and achieve the results needed in the terminal state exams in order to progress to further education and training opportunities. Once learners had acquired basic communication skills, the development of the reading and writing skills needed for successful learning across a wide range of subject areas became a priority. In this context it was arguably not appropriate to ask learners to take time out from the regular work of the programme to record interviews and conversations where the activity's primary purpose was to provide data for research. Recordings could have been made outside of school hours, but as their language support teacher this researcher decided that any time spent with the students should be unambiguously focused on supporting learning rather than on doing research. On the other hand the writing process mentioned above was a standard component of the programme, designed to provide valuable learning opportunities, and was regularly used by all learners.

Collecting language samples which are produced as part of a regular programme of work may also help to reduce or minimize the ' observer effect' (also referred to as the Hawthorne Effect). Researchers have found that participants' performance can be affected when they know they are being studied, and Mellow et al. (1996, 334) suggest that this observer

effect is of particular importance for SLA research as it 'may be the single most serious threat to studies of spontaneous language use'.

Working with written data has some advantages. Oral recordings are more difficult to transcribe accurately and some detail – such as the suppliance of auxiliaries and suffixes– may be lost due to elision etc. Written data can also result in longer more elaborate text than conversation. Bardovi-Harlig (2013) provides specific examples of this, in a comparison of data from conversational interviews and written journals. She found that learners gave much more detailed accounts of the same topic in their journals than in responses to questions during interviews. Bardovi-Harlig comments that when the learners write 'they must set the scene since a reader cannot contribute to the construction of the text in the same way as an interlocutor' (237-238). It has also been suggested that written data may provide more information on the outer limits of the learner's current linguistic knowledge. According to Saville-Troike (2006, 164) writing can potentially push learners to these outer limits 'because writers must express ideas without recourse to objects and events in their own immediate physical environment or that of their reader(s), or to gestures and other nonverbal means of communication, and without reliance on immediate feedback or hearer cooperation to fill in gaps...'. Writing is, she suggests, 'probably the most dependent of the four language activities on linguistic knowledge'.

4.5 The Research & Learning Context

4.5.1 Introduction

The learners involved in this study were students in an Irish post-primary school, and the data was collected in the course of a writing process developed for, and routinely used in, the school's Language Support Programme (LSP). This section provides a short introduction to the school, the programme and finally, the writing process.

4.5.2 The school

St John's[1] is a large Dublin community school, much of whose catchment area has been described as 'disadvantaged' on the basis of socio-economic indicators such as unemployment levels, housing conditions and educational attainment[2] St John's has an exceptionally large number of students from L1 backgrounds other than English - almost 200, or 25% of the total school population. In comparison a recent Economic and Social Research Institute (ESRI) study estimates that in the post-primary sector as a whole 'newcomers[3] make up a relatively modest proportion of students, typically 2 to 9 per cent, within each school' (Smyth et al.,2009, XIV) In all over 35 languages are spoken in St John's, the most widespread being Polish, Lithuanian and Malayalam (the L1s of the three learners in this study) Romanian, Somali, Yoruba and Lingala.

4.5.3 The Language Support Programme

During the period of this study the school's Language Support Programme (LSP) catered for between 70 and 100 students every year, most of whom would have arrived in Ireland within the previous 24 months, and many of whom would have had little or no English at the time of arrival. The Department of Education's allocation of 44 teaching hours a week was used to timetable ESL classes for students from six different year groups and over 25 different class groups, with an age range of 12 to 19 years. Because of the constraints imposed by the mainstream timetable, it was often not possible to group 'like with like' and an ESL teacher sometimes worked with several different age groups and acquisition levels during the same 35 minute lesson period. The three learners in this study received, on average, 7 periods of language support a

[1] Not the school's real name
[2] based on data in the 2006 Census of Population
[3] Newcomer is the official term for students from immigrant families and it includes English speaking immigrants from the UK, USA, South Africa etc as well as students whose L1 is not English.

week, usually as part of a group of 8-12 students from the same age cohort.

The general goal of the LSP is to enable ESL students acquire the linguistic competence and cultural knowledge they need to become active members of Irish society. An important specific task of the programme is that of helping students access the educational opportunities offered by the school, therefore particular attention is paid to the language interaction of the mainstream classroom and the language demands of independent study.

In the programme there is an emphasis on the active use of language for communication, rather than the study of language as a rule governed system (Littlewood, 1981; Widdowson, 1978). In relation to grammar, form-focused instruction is provided since research suggests that this supports successful second language acquisition (Spada and Lightbown, 2008). However as far as possible this focus on form is fully contextualised within communicative activities, occurring incidentally as individual students are assisted in comprehending or conveying more effectively the meanings they are currently engaged with, in this way providing them with 'psycholinguistic data optimized for acquisition' (Ellis, N., 2008a, 122). An example of such form focused instruction is provided in Section 4.5.4.

Throughout the programme extensive use is made of authentic texts in the belief that they provide the kind of input which is most likely to activate natural language learning processes (Little, Devitt and Singleton, 1989) and because they facilitate the promotion of learner autonomy in the language classroom (McGarry, 1995). The LSP policy in St John's places considerable importance on the development of learner autonomy. Over the last 20 years the achievement of learner autonomy has increasingly been seen as a central goal of education and it is now an aspiration enshrined in many syllabuses and curriculums:

> Students who are encouraged to take responsibility for their own work, by being given some control over what, how and when they learn are more likely to become actively involved in the learning process ... they are more likely to be able to set realistic goals, plan programmes of work, develop strategies for coping with new and unforeseen circumstances, evaluate and assess their own work and generally to learn how to learn from their own successes and failures in ways which will help them to be more efficient learners in the future. (McGarry 1995, 1).

Successful *language* learning, it has been argued, depends crucially on students achieving a substantial degree of autonomy as language users: 'Communicative efficiency in the target language community depends on learners having the independence, self-reliance and self-confidence to fulfil the variety of social, psychological and discourse roles in which they are cast' (Little 1991, 27).

Students in St John's LSP are encouraged to become actively involved in setting goals, planning work programmes and monitoring progress. In practice this means, for example, that classroom activities and tasks are negotiated with the teacher, or chosen by the individual students from a range of options. In the writing programme students are shown how to use checklists and other aids to help them become increasingly independent as assessors and editors of their own work.

The challenges faced by ESL students in the Irish post primary system are considerable, not least because of the limited time and resources allocated to language support programmes (Fionda 2010; Lyons and Little 2009). To successfully meet these challenges, it can be argued, the achievement of a significant degree of learner autonomy is not just desirable but essential. Research suggests that it can take 5-7 years to acquire the cognitive/academic language proficiency (CALP) needed to successfully access a school curriculum through a second language (Collier 1987, Cummins 1981). Therefore the terminal state examination, the Leaving Certificate (LC) , with its requirement of pass grades in at least 5 subjects,

represents a major challenge for the students in St John's LSP, most of whom will have been in Ireland less than five years when they take the exam.

For a variety of reasons, ESL students can regard success in the LC English exam as a key goal of their learning, and yet this is probably the most difficult exam for them to succeed in. At both Higher and Ordinary Level, the English exam is divided into two papers which between them require the candidate to write for approximately 6 hours, producing a variety of different texts, some over 1,000 words long. Fluency and grammatical accuracy are taken into account in the marking of all tasks under two criteria (Efficiency of Language Use and Accuracy of Mechanics) which together constitute 40% of the total marks.

In St John's the LSP employs a number of strategies specifically aimed at enabling students achieve the writing skills needed for success in LC English. One of these is a process designed, among other things, to help students develop the ability to write confidently and fluently in time constrained conditions. This is the writing process from which the data for this study has been drawn and it is described in detail in the next section.

4.5.4 The writing process

The range of tasks and varieties of approaches to writing employed in the LSP reflect the kinds of written work ESL students are asked to undertake in mainstream subject learning and the different conditions under which they are expected to perform, from guided classroom activities to independent homework to exams. While some of the work in the LSP classroom is carefully planned by the students, with exemplars, templates and other supports provided by the teacher, there is also a writing process designed to encourage the students to write as spontaneously as possible. The process employs the kind of 'brainstorming' elicitation methods

regularly found in creative writing programmes, however it is used in the LSP as a starting point for a wide range of different writing tasks, from responding to a poem or expressing an opinion to relating a personal narrative.

The aim is to produce a 'burst' of writing – sometimes as an end in itself, and sometimes as a draft or preparation for a more carefully planned piece of work. The student is usually allowed just 10 or 15 minutes and encouraged to write as much as they can in the time. However, as their proficiency develops and their capacity for sustained writing increases, the learners may choose to spend longer on these writing tasks, and some of the texts sampled in the later stages of this study were written in 20 or 25 minutes. To avoid interrupting the flow of ideas the student is told to use an L1 word if they cannot immediately think of an appropriate English word. Where the student is continuing on with a narrative already begun, or making a second attempt at a task, or returning to a topic dealt with previously, they may read over the earlier texts, but these are then put aside before the writing of the new text commences. Several photocopies of the student's text are made and kept in a folder for further work on content, style and grammar.

The clear separation between the initial 'creative' process and the subsequent 'editing and correcting' process is designed to help encourage the students to write relatively spontaneously, to focus on meaning rather than form.

Because these are relatively spontaneously written texts they can provide both teacher and learner with insights into the current stage of development of the learner's implicit knowledge of English. And when the photocopies of these texts are subsequently used for form focussed instruction, for example, the teacher may provide feedback on the learner's developing language, drawing attention to particular features of grammar and helping the learner to construct checklists appropriate to

their current stage of acquisition, checklists which the learners can then use with other writing assignments to predict and identify errors –thereby becoming more independent as editors and correctors of their own work.

4.6 The learners

4.6.1 Introduction

This researcher had obtained the permission of the Board of Management of St John's to collect data from ESL students for a longitudinal study of SLA acquisition. For this study of the later stages of language development, informed consent was sought and obtained from three learners entering the second year of the Junior Cycle, and informed written consent was obtained from their parents (when the learners reached the age of 18 they reaffirmed their consent in writing).

These three learners were chosen both because of their similarities and their differences:

(a) they were in the same ESL group and had this researcher as their language support teacher.

(b) they were also in the same mainstream Junior Cycle class and were together in the same lessons with the same teachers for 26 out of 28 school hours every week.

(c) in conversations and in a learner profile completed at the end of First Year, they had reported a similar pattern of exposure to English outside the school classroom.

(d) In terms of overall language development, their rates of progress appeared to be different, relative to their respective starting points at the beginning of First Year; at the end of First Year however all three were, broadly speaking, at a similar stage in the development of tense-aspect morphology – with, for example, rates of appropriate use of the Simple Past tense at around 70% - 80%.

The three learners chosen for this study have L1s which differ from each other in many respects, one belonging to the Dravidian family of languages and the other two to different branches of the Indo-European family. Malayalam, spoken by approximately 38 million people, is a South Dravidian language whose closest relative is Tamil. It is an agglutinative language with its own Brahmic writing system based on a syllabic alphabet (Asher and Kumari, 1997). Lithuanian, spoken by about 3.5 million people, is one of only two living Baltic languages (the other being Latvian); Polish, spoken by approximately 50 million people worldwide, belongs to the Lechitic group of West Slavonic languages. Both Polish and Lithuanian are strongly inflecting-fusional languages and both use the Latin alphabet with diacritics (Balode and Holvoet, 2001; Stone, 2009).

The Learners chose their own pseudonyms for the study – the Lithuanian student chose 'Ana', the Polish student 'Luisa' and the Indian student, whose L1 is Malayalam, chose 'Sara'.

4.6.2 learner profiles

Ana and Luisa had both arrived in Ireland at the age of 13, just weeks before starting First Year in St John's. Both had completed their primary education through their L1, and while their primary schools had offered some lessons in English as a foreign language, both were at a relatively early stage of English language development. Initial assessments showed that, in terms of tense-aspect morphology, both were heavily dependent on the copula *be*, base forms of the verb and the *–ing* participle. Luisa displayed a wider vocabulary and greater control of syntax, however, and subsequently, over the course of First Year, appeared to make quicker progress than Ana.

Sara was 12 years old on enrolling in St John's. Most of her primary education had been through Malayalam, however she arrived in Ireland a

year earlier than Ana and Luisa, and spent three terms in an Irish primary school before starting First Year in St John's. As a result her English language development was more advanced than the other two. Initial assessments, for example, showed productive use of the Present and Past Progressives and of the Simple Past (with irregular verbs). By the start of Second Year, and the commencement of this study, Luisa was displaying greater lexical proficiency than Sara and had narrowed the gap between them in terms of morphological development and overall accuracy of language use.

All three learners report that in their first three years in St John's (including years 1 and 2 of this study) they mainly socialised, outside the school classroom, with teenagers who spoke their L1 - this was easy to do as there were dozens of speakers of their L1s in the school and more in the local community. To a lesser extent they also socialised with other L2 English speakers they had got to know through the language support programme. They spoke their L1 at home and used Skype and other internet means to communicate with friends and family abroad – Sara did so very frequently, almost daily, and Ana reports spending at least two hours a week on Skype during her first year in Ireland. While they occasionally went to the cinema with their friends to watch English language films, at home they mainly used the internet or satellite television to watch films and programmes in their L1 – again, Sara did so to a greater extent than the other two. All three agree that a lot of their exposure to L1 English took place in the classroom, or during extra-curricular school activities such as sport (Ana and Luisa were on the school athletics team in Second and Third Year).

For Ana and Luisa the pattern of exposure to English changed somewhat during their final two years at school (years 3 and 4 of this study) with both forming close friendships outside their L1 group and increasingly socialising through English – whether with other L2 English speakers or native speakers . Ana, for example, reports that during this period she

mainly texted in English and watched DVDs and television programmes mainly in English. Throughout the period of this study Sara's closest friends were Malayalam speakers, and she continued to socialise, both inside and outside school, mainly within the Malayalam-speaking community.

Of the three learners, Luisa had the strongest literacy skills in her first language. She attended weekly Polish lessons throughout her time in St John's and her Polish teacher considered her to be an exceptionally good writer, and an enthusiastic reader who was prepared to engage with demanding texts. Luisa achieved an A grade in L1 Polish[4] in the Leaving Certificate. Ana spoke some Polish and Russian (which she had learnt from family members whose L1s these were) in addition to Lithuanian. However she reported that in primary school Lithuanian and other languages had not been among her stronger subjects, that she preferred Maths and Science for example. She achieved a C grade in L1 Lithuanian in the Leaving Certificate. Sara reported that she wrote infrequently in Malayalam once she came to Ireland, apart from short emails and instant messaging texts to friends in India. She enjoyed reading and read mainly popular fiction in English.

All three learners ultimately took Higher Level English in the Leaving Certificate, achieving C grades, and went on to third-level education in Ireland, pursuing courses in computer studies or business and administration.

5 Data Collection, Transcription & Coding

5.1 Data collection

[4] In addition to the European languages offered as part of the Leaving Certificate curriculum (e.g. French, Spanish, German, Italian, Russian) EU citizens may sit a native-speaker level exam in their L1 as part of the Leaving Certificate. These languages are referred to as 'non-curricular' by the State Examinations Commission, because they are not taught in schools.

For the reasons outlined in Section 4.4.3 above, the aim was to collect written data which was as spontaneously produced as possible. To this end all the language samples used in this study were produced using the writing process described in Section 4.5.4, a writing process which the learners had become familiar with during the year before the commencement of the study. While the learners were, of course, aware that samples of their work were being collected for research purposes, they showed no signs of being particularly conscious at the time of writing a text that it might later be included in the data – perhaps because the writing tasks were undertaken in the context of routine classroom activities stretching over a period of four years.

The learners were always given the option of not doing a proposed task, or of doing it at a different time – this was the general practice observed in the classroom, part of an overall approach designed to foster learner autonomy (see Section 4.5.3). The learners were also free to suggest an alternative topic or task, and some of the personal narratives and creative writing samples included in the data were the result of tasks set by the individual learner herself. In terms of sampling natural language use it could be considered a strength of this data collection process that the samples were not taken from a regularly scheduled task – done say every Wednesday morning at 10am – but rather that the samples were taken from texts produced when the learners were in the mood to write and had chosen to write.

The writing process from which the data was drawn was just one of a number of approaches to writing employed in the programme, and the frequency with which it was used depended on a number of factors, the most important of which was the work which the learners were then undertaking in the mainstream English class – because a primary aim of the language support programme was to help students access the curriculum and successfully prepare for the state exams at the end of the Junior and Senior Cycles. Sometimes the main focus of a week's work was

on carefully planned, constructed and edited written texts, rather than on spontaneous writing, and sometimes when the learners were being asked to tackle challenging literary texts in the mainstream class - a Shakespeare play for example - the focus might be on reading and discussion, rather than writing. The frequency and spacing of data collection was also affected by the school calendar, as is often the case with longitudinal research conducted in educational institutions. As a result, while all samples of spontaneous writing produced by the three learners were included in this study, the samples were sometimes only separated by a day or two, but generally by a week or two. And no samples were collected during the school holidays. In the ESF studies of temporality the average period between two sampling points was one month and Klein expressed the opinion that because the developmental process was so slow, the risk of missing important developments was 'almost non existent' (1995a, 31).

In terms of the kind and range of language sampled, drawing the data from a writing process which samples the themes and tasks of the mainstream English classroom did not in itself impose many limitations. The English curriculum in both Junior and Senior cycles requires students to interact with a wide range of texts and to themselves produce a wide variety of written texts. For the Leaving Certificate for example, the syllabus specifies that the students must be proficient in the narrative, persuasive, argumentative, descriptive and aesthetic uses of language. Typical writing tasks sampled in the data include expressing hopes and opinions, giving advice, responding to poetry and prose texts, recounting past experiences, summarising the plots of novels, films and plays, and the writing of fictional stories.

As a result the data includes both nonnarrative monologic discourse (description, argument, irrealis) and narrative extended monologic discourse (personal and impersonal narratives). As Bardovi-Harlig (2013, 248) points out, narrative texts allow researchers to document the learner's use of tense aspect in a past-time chronologically ordered

framework: 'The foreground will exhibit perfective past (simple past in English). In contrast, anterior, simultaneous, and future events and situations in the background all provide contexts for the use of diverse tense-aspect morphology. These factors combine to create a rich environment for the study of temporal expression.' On the other hand, Bardovi-Harlig argues that the inclusion of a variety of nonnarrative texts enriches L2 tense-aspect research by providing environments for the expression of temporal relations other than forward movement of time. Noting that nonnarrative discourse has not been as widely investigated as narrative has, Bardovi-Harlig suggests that this kind of data needs to be more fully exploited by researchers.

While some tasks could be viewed as challenging and others as comparatively easy, it appeared that, in general, the learners struggled to express themselves, when they did, mainly because of the meanings they had chosen to express, or the means of expression they had selected, rather than because of an inherent difficulty in the theme or task. For example, when the learners were asked to give advice on how to prepare for exams, Ana and Luisa confined themselves, in the main, to straightforward suggestions in simple, emphatic English : '*Never revise the day before the exam. The best idea is to go out with your friends and get some fresh air*' (Ana, file 83). '*You have to focus and listen to your teachers and do exactly what you're told*' (Luisa file 96) . Sara, made similar suggestions but chose more complex ways of expressing them and occasionally struggled with the syntax or grammar, for example: '*Make sure before you went into the exam you [/] everything you've studied is in your long term memory. If its not your hard work <will be waste> [/] is [/] was a waste*' (Sara file 78, [/] = preceding word/phrase deleted).

It is also the case that while some topics and tasks might seem likely to elicit particular tense-aspect forms, learners have the scope, in free writing, to approach the task or topic in ways which create contexts for other forms. For example, when asked to predict how a story was going to end, Ana responded to the question ' What is Dazza going to do next?'

with a present tense narrative, while Sara and Luisa both employed future forms. And given the topic 'My Ideal Home' Ana used mainly lexically future verbs like *'want' 'have to'* and *'like'*, while Sara relied predominantly on the Will Future and Luisa mainly used modals like '*would*' and '*should*'.

In any case, for this study no attempt was made to manage the sampling in a way which would provide, for example, a minimum number of obligatory contexts for particular forms. Also, the setting up of the writing tasks did not include the pre-teaching of specific vocabulary or a focus on specific features of grammar. Any class time spent preparing for a task generally consisted of a short 'brainstorming' session on the relevant topics or themes.

As a result there are forms which are likely to be infrequent in the data – arguably because they are infrequent in the naturally occurring language use characteristic of writing activities undertaken in mainstream English classrooms. Interrogatives, for example, are likely to be infrequent in such written data although some occasions for their use may arise in fictional and personal story-telling (in dialogues and reported speech), letters, emails etc. However, because of the longitudinal nature of this study, its data does manage to capture the various stages the learners pass through on the way to acquiring target-like use of the range of question forms. Similarly the written tasks undertaken in the mainstream classroom would appear to provide infrequent opportunities for the use of future forms, however once again, while the *percentage* of future contexts in the data is low (5% of all tense-mood-aspect contexts) the *number* of contexts (468) is large enough to provide insights into the learners' acquisition of future expression.

When writing samples were photocopied for future classroom work (see Section 4.5.4) additional photocopies were made at the same time and set

aside for later transcription, coding and analysis using the CHILDES suite of programmes.

5.2 Transcription

5.2.1 Introduction

The invention of the tape recorder transformed language research in the second half of the twentieth century, making it possible to collect large amounts of raw data from a variety of subjects in a wide variety of contexts. But the spoken word still had to be painstakingly transcribed, and analysis involved long hours spent scanning pages, pencil in hand. When Roger Brown wanted to share his data with others, the transcripts had to be typed onto stencils so that multiple copies could be run off. In the nineteen eighties, however, the introduction of the powerful and affordable personal computer opened up the possibility of databases and analytical programs; and the development of the internet provided opportunities for collaboration unimaginable only a few years before.
Looking back on this period Brian MacWhinney wrote:

> [T]he possibility of utilizing shared transcription formats, shared codes and shared analysis programs shone only as a faint glimmer on the horizon, against the fog and gloom of handwritten tallies, fuzzy dittos, and idiosyncratic coding schemes. Slowly, against this backdrop, the idea of a computerized data exchange system began to emerge. (2010, 8)

From a tentative beginning in 1981, MacWhinney and Catherine Snow went on to develop just such a system for the study of first language acquisition. Drawing on the expertise of a large number of researchers, the Child Language Data Exchange System (CHILDES) database was established, along with a transcription format (CHAT) and an analysis programme (CLAN). The first set of manuals was published in 1991.
Today, CHILDES is the world's largest corpus of spoken language data, containing over 44 million words from 28 different languages. As well as

data collected over the last two decades it also includes earlier corpora – such as Brown's transcripts of Adam, Eve and Sarah – which have been reformatted into CHAT. The extent to which the system has facilitated research can be judged from the fact that a 2002 review identified more than 2000 articles which had used the CHILDES data or the programs (MacWhinney, 2008b).

The language samples which provide the data for this study have been transcribed by the researcher into CHAT format and analysed with the help of the CLAN programs. The resulting corpus consists of over 72,000 words, of which 64,500 have been fully analyzed (the remainder, comprising words subsequently deleted in the course of writing, have been coded on the main line and included in the analysis of error corrections and changes (see Section 6.3).

5.2.2 Adapting CHAT for written data

The CHILDES project is principally concerned with research into first language acquisition and the CHAT format was designed and developed for the transcription of recordings of face-to-face conversational interactions. However CHAT has also been used to transcribe written discourse, and the manual does give some specific symbols for this, mostly for orthography and punctuation errors. For the purposes of this study it was found necessary to make some adaptations – reassigning existing symbols to new uses, or creating new symbols. This section explains the main decisions made in this regard, in the process discussing some of the problem areas for the transcription of written data.

Distinguishing between transcription and coding, MacWhinney (2010, 17) says: '[T]ranscription focuses on the production of a written record that can lead us to understand, albeit only vaguely, the flow of the original interaction'. In relation to this study it was considered important that the transcription of the written data provide some sense of the process as well

as the product. Especially when it is produced relatively spontaneously, a written sentence may undergo a number of changes during its production– insertions, deletions, corrections, redrafting – and these can be of interest to the researcher, especially in terms of the degree to which a form has become established in the interlanguage – whether, for example, in a given instance, it is supplied appropriately and accurately initially, or only subsequently as an afterthought. An analysis of the changes and self-corrections made in the course of writing can also provide some indication of the learner's orientation to the task – of the degree to which they were focused on the communication of meaning rather than the achievement of accuracy.

In conversation a speaker may start to say something then stop, repeating the basic phrase but changing the morphosyntax while maintaining the same idea: ' *the fish is..... the fish are swimming*'. In CHAT format the symbol [//] is used for these 'retracings with correction' and MacWhinney (2010) recommends that with written data, this symbol should be used where phrases are deleted and rewritten, with the symbol [: 0] used where single words are deleted. Written data provides the researcher with the possibility of making a further distinction which could give a more nuanced picture of the learner's knowledge of a form – because in addition to the *immediate* change, where a word or series of words are deleted and substituted during the initial process of constructing the phrase, there are *later* changes made after the phrase or sentence has been written, but before the writing task itself has been completed – in other words where the learner has reviewed the construction, moments or minutes afterwards, and made a judgement about its acceptability. These later changes are relatively easy to identify because they typically necessitate insertions. For this study it was decided to use [//] for later changes and [/] for immediate ones. [/] is normally used in CHAT for 'retracing without correction', in other words for when a speaker immediately repeats the same word or phrase. This rarely occurs in the written data here and when it does – '*the the dog*' for example – it is

treated as an error. (For other symbols adapted or introduced for this study, see Appendix 1)

MacWhinney (2010) talks about the problems of producing transcriptions in a way which meets the needs of two very different audiences : the human audience of transcribers, readers and analysts on the one hand, and the digital computer and its programs on the other. In the case of written data it is generally possible to achieve an accurate and readable mainline representation of what happened. However where certain later changes or insertions are made, careful thought has to be given to the best way of representing these. Transcription difficulties can occasionally arise for example, where changes are made simultaneously at several different places in a sentence; or where a later edit or correction requires a change earlier in the sentence and this is not made, creating a consequent error in the final version of the sentence. In both cases phrases need to be scoped carefully for coding in order to give the reader a clear picture of the initial construction and the process by which it was changed.

5.2.3 Transcribing the data

The written samples were transcribed into CHAT format by this researcher, creating files which consist of (a) 'headers' providing information about the learner and the particular writing context, and (b) the main tiers or 'main lines' , which record, sentence by sentence, the text the learner wrote. These 'base' files were proofed by sight against the original handwritten texts twice, and once again using the CLAN mor xl command (which can be used to identify typographical errors as it lists all the unrecognised words in a file). The files were proofed again against the originals during the next stage of the process, when errors were coded on the main tier. Subsequent morphosyntactic coding also involved incidental checks on the accuracy of the main tiers. CHECK commands were run after every alteration or addition to ensure that the files were correctly formatted and that the analytical programs could access the

data. The final versions of the transcript files contain two dependent coding tiers in addition to the main tier. The coding process is described in detail in the next section.

5.3 Coding

5.3.1 Introduction

In order to investigate the research questions which guide this study a complete morphosyntactic analysis of all the data was required, and is found in the transcript files on a dependent %mor coding tier located just below each main tier. It was also necessary to identify the tense-mood-aspect contexts created by the learners, so that these could be compared with the forms actually supplied in each context. These are coded on a dependent %tar tier located below each %mor tier. As part of the process of acquiring a picture of the learners' overall language development, all learner errors were coded on the main line. Information on other areas of development, such as lexical diversity, was retrieved by running analytical programs on the %mor tier.

So a sentence written by Ana, for example, when transcribed and coded for this study, takes the following format:

*ANA: when I first rode oa bicycle [^c] I was about six years old .
%mor: conj:subor|when pro|I adv|first v|ride&PAST odet|a n|bicycle pro|I v|be&PAST&13S prep|about det:num|six n|year-PL adj|old .
%tar: $POS:PST:IRR [^c] $POS:PST:COP

Appendix 1 contains a full list of codes and conventions.

5.3.2 Error coding

In CHAT convention, detailed comments and codes for errors can be placed on a dependent coding tier (%err), with the errors also marked with the symbol [*] on the main line. However as MacWhinney (2010) notes, for most work it is not necessary to use the dependent tier, and in fact supplementing [*] with additional codes on the main line actually facilitates easier retrieval of error types. For this reason, main line error coding is used in this study. Misspellings where the spelling was phonetically accurate or approximate, to the extent that the word when read aloud would sound target-like (*suprise* for *surprise*, *excercise* for *exercise* etc) were not treated as errors. Instead a CHAT replacement procedure was used so that the target word appears on the %mor tier and is correctly identified by the analytical programs. The same procedure was used for cases where the possessive apostrophe had been omitted : *my brothers [: brother's] car*. A full list of the error codes used in this study is given in Appendix 1.

5.3.3 Morphosyntax analysis

A major advantage of using the CHAT format is that it is then possible to run the CLAN program MOR on the transcript files. This is a very useful analytical tool as it can be used to generate a coding tier for every sentence, each word being labelled by syntactic category, followed by the word itself broken down into its constituent morphemes. When it is run, the program provides a comprehensive list of categories and morphological analyses for each word. For the word *that*, for example, it will offer a number of options – determiner, relative pronoun, demonstrative pronoun, intensifying adverb etc . To select the correct form for each ambiguous case the file can be edited 'manually' using Disambiguate mode, or automatically using the POST program developed by Christophe Parisse (who also developed POSTTRAIN which researchers can use to improve POST's performance on their own files).
It has been suggested by MacWhinney (2008b) that few, if any, human coders can achieve 95% accuracy on a 'first pass' through data. Careful

proofing (and reproofing) is clearly necessary if the aim is to achieve at least this level of accuracy. For this study, the time taken to carefully proof and correct an automatically disambiguated POST file, was compared with the time taken to manually disambiguate, then proof and correct the same file, both versions containing identical coding in the end. Initially 20 files were processed using both procedures and the time taken and the number of changes made at the proofing stage were noted. While corrections took longer for POST (with 10%-15% of the coding being changed on average) overall it was marginally quicker, and this method was used for the remaining files. Each POST file was proofed three times and where any difficulty arose about choice of category, Leech and Svartvik (1994) was consulted.

MacWhinney points out that main line transcription has to serve two, sometimes conflicting functions. On the one hand it has to represent the form of the speech as actually produced, on the other hand it must provide input for morphosyntactical analysis. Difficulties arise, for example, when the learner produces a malformed word. The Clan Manual suggests using replacement **[:** *target word***]** on the main line so that MOR can correctly identify the word. For this study of written data it was decided, for clarity and consistency, to use replacement only for misspellings, and to use a different approach for other malformed words.

In the following sentence, for example, *felt* is not a spelling error - at this stage of her L2 development, Ana is using *'t'* inappropriately as a past tense inflection for a number of verbs. For this reason *felt* is marked as an error on the mainline:

ANA: I'd started to panic so I aga [/] felt [fell] again .

However the MOR program will then analyze *felt* as a noun: n|felt or as a different verb: v|feel&PAST. It was decided to manually edit the %mor tier in such cases, supplying the correct target form but using an asterisk to indicate that an error was made: *v|fall&PAST. This has the advantage that results generated by program searches on the %mor tier

(for example a FREQ list of Simple Past verbs) will clearly indicate that the correct form was not supplied in this instance. The same approach has been used where learners produce a nonnative form based on their L1 (like *manages* for the plural noun *brakes*).

The MOR program is triggered by the use of the [/] and [//] symbols (discussed in 5.2.2 above) to ignore all deleted or 'retraced' material in line with the CHAT convention that it is the final version which is morphosyntactically analyzed as the sample of language. For this study it can be of interest whether a learner produced an appropriate or accurate version of a form on a first or subsequent attempt and that information can be retrieved from the main line tier using other programs such as KWAL or COMBO.

5.3.4 Context coding

The tense-aspect contexts created by the learners in the course of their writing were identified and coded on a second dependent tier (%tar). Each finite verb clause was coded according to the form which would be preferred by an L1 speaker in that context : positive, negative or interrogative; past, present, future; simple, progressive or perfect etc as appropriate (the full list of codes is given in Appendix 1). Where there was any doubt about the target tense-aspect form – because the meaning as expressed by the learner was ambiguous or difficult to decipher, for example – the clause was excluded from the analysis. In practice only a small number of clauses had to be excluded on this basis (less than 1%).
The information contained on this coding tier was used to identify the emergence of contexts for particular tense-aspect forms, as well as to calculate rates of appropriate and accurate use of tense-aspect forms, in addition to other calculations such as the relative frequency of past, present, and future time contexts in the data.

5.4 Analysis

5.4.1 Lexical development

Lexical diversity, the range and variety of vocabulary deployed by a speaker or writer, is considered a useful predictor of overall language proficiency (Jarvis, 2013) and lexical diversity indices have been used for many years in first and second language research (McCarthy and Jarvis, 2007). Most indices of lexical diversity have been based on the ratio of different words (types) to the total number of words (tokens). However Type-Token Ratio (TTR) and mathematical transformations of the TTR are subject to sample-size effects. Samples containing larger numbers of tokens give lower values for TTR (McKee et al., 2000).

For this research, lexical diversity has been measured using an instrument designed to overcome the problem of sample size. The *vocd* program, developed at the University of Reading (Malvern and Richards, 2002), is based on an analysis of the probability of new vocabulary being introduced into longer and longer samples of speech or writing. This yields a mathematical model of how TTR varies with token size. By comparing the mathematical model with empirical data in a transcript it provides a measure of lexical diversity referred to as *D*. A recent study (Crossley et al., 2011) which compared a variety of lexical indices (measuring variables such as content word frequency, polysemy, hypernymy,and word length, as well as diversity) found that D was the best predictor of human ratings of lexical proficiency.

In an evaluation of *vocd* McCarthy and Jarvis (2007) suggest a number of stable ranges within which texts of varying lengths can be compared, and in line with their findings, for this research only texts in the range 100-400 tokens have been included in the analysis of lexical diversity. The analysis was conducted on the morphosyntactically coded tier –as a result false starts and deletions were not counted – and all word tokens were lemmatized. The data was divided into 8 time periods, (two per school

year, each covering approximately four months) and mean D scores were calculated for each period.

The information on lexical diversity- which essentially reflects the rate of word repetition - is supplemented with an analysis of word types and frequencies in the learners' data. For this, statistical information was extracted from the data with the aid of the CLAN FREQ program and transposed to EXCEL where further calculations were made. The CLAN KWAL and COMBO programs were used to identify and extract specific examples from the learners' coded transcripts.

5.4.2 Accuracy of language use

For this study the measure chosen for accuracy of language use is *errors per 100 words* (the number of errors divided by the total number of words produced, multiplied by 100). As Ellis and Barkhuizen (2005) note, it is a general measure of accuracy which has been widely used. A more fine-grained analysis is also provided by comparing the relative frequencies of different types of errors in the learners' interlanguage. All errors (excluding spelling, punctuation and other orthographic errors) were coded on the main (speaker) tier of each transcript, then information on the frequencies of different errors, as well as total error counts, was extracted using the CLAN FREQ program and transferred to EXCEL spreadsheets where the data was divided into 8 time periods (two per school year) and the types and rates of errors per period were calculated.

Errors initially made by the learner but then corrected, either immediately or subsequently in the course of producing the text, were excluded from the *error per 100 words* analysis. These were however included in a separate analysis which looked at the learner's orientation to the writing task (whether the main focus was on meaning or on formal accuracy).
For this analysis, a comparison was made between the percentage of text changes which involved changes in meaning and the percentage where the change was made with the clear intention of correcting an error (or an

assumed error). All instances of false starts, deletions, rephrasings and self-corrections, coded on the main speaker tier with the symbols [/] and [//], were identified and quantified by time period using FREQ and the information transferred to EXCEL where further calculations were made, including the number of changes ([/]+[//]) per 100 clauses, and the percentage of changes which involved error self-corrections.

5.4.3 The acquisition of temporal expression

Past-Related Forms

For the analysis of the development of the learners' past-related temporal expression, it was necessary to establish rates of appropriate use and accurate use for the Simple Past. It was also necessary to identify and record overgeneralisations – both of other forms in Simple Past contexts, and of Simple Past forms used in other tense-aspect contexts. With the aid of the CLAN programs FREQ, COMBO and KWAL, all verb forms used in contexts coded as Simple Past on the %tar tier were recorded on a database by file and period. Overregularised forms such as *runned, flyed*, etc were counted as appropriate use, but not as accurate use. If another past form, the progressive for example, was used in a Simple Past context, then the overgeneralisation of this form was recorded both in the Simple Past data and in the data on that form. Rates of appropriate use were calculated as the ratio of the number of appropriate forms to the number of obligatory contexts. Accuracy rates were calculated in the same way. All rates are given as percentages.

Negative statements and questions were not included in this Simple Past analysis as in these constructions tense is generally marked on the auxiliary not on the main verb . These were however included in a separate analysis (6.3) of the development of negation and interrogatives.

The use and overuse of the Present Perfect and Past Progressive were established in the same way as for the Simple Past. To be considered an

appropriate use in a Present Perfect context, a form had to consist of the auxiliary *have* and a main verb, but not necessarily the perfect participle; for the Past Progressive however both the past tense auxiliary *be* and the *–ing* participle had to be employed. This is in line with the practice adopted by Bardovi Harlig (2000). As this analysis concerns tense-aspect only, agreement of person and number were not taken into account because agreement belongs to a different semantic system – as a result *he have lived here since May* would be considered both an appropriate and an accurate use of the Present Perfect. Following Bardovi-Harlig (2000) emergence of the Present Perfect was defined as the appropriate use of the form with three distinct lexical verbs.

Verbs were assigned to lexical aspect classes - state, activity, event (accomplishment and achievement) – with the aid of the tests devised by Shirai and Andersen (1995), e.g:

Step 1: State or nonstate
Does it have a habitual interpretation in Simple present?
If no - State (e.g., *I love you*)
If yes - Nonstate (e.g., *I eat bread*) Go to step 2
Step 2: Activity or nonactivity
Does 'X is V-*ing*' entail 'X has V-*ed*' without an iterative/habitual meaning? In other words, if you stop in the middle of V-*ing*, have you done the act of V?
If yes - Activity (e.g., *run*)
If no - Nonactivity (e.g., *run a mile*)

In the learners' L1s (Polish, Lithuanian, Malayalam) the tense of the original expression or thought is usually retained following reporting verbs, such as verbs of speaking, cognition, emotion, and perception. It can take time for learners from these L1 backgrounds to recognise that backshift is often required in English and initially none of the learners in this study employed backshift. Instances of present forms being supplied instead of the past forms in such constructions (reported speech etc) have been excluded from the analysis as, given the learners' stage of morphological development (see Section 7.2.2) , these instances most likely represent a failure to backshift rather than a failure to supply the appropriate form.

Future Forms

For the analysis of the development of future time expression all verb forms used in contexts coded as Future on the %tar tier were recorded on a database by file, period and text type (e.g. personal narrative, fictional narrative, description, argument or opinion piece, response to a text etc) . Verb forms found with future-in-the-past reference or in backshifted future contexts with reporting verbs were also included in the analysis, because it is always clear which future form has been used. In the following examples (a) is counted as an instance where the learner used the Will-Future, and (b) as an instance where the learner used the Go-Future:

(a) *She said she would come at nine.*

(d) *We promised ourselves that we were going to live in this house.*

The meaning associated with each verb form found in a future context was also recorded by type of future (based on Leech, 2004; Radden and Dirven, 2007), eg: Predicted Future (*Sarah will keep her promise*) ; Intentional Future/Future of Present Intention (*I'm going to learn to fly*); Contingent Future/Future of Present Cause (*it's going to rain*); Planned Future (*I'm getting married next month*); Scheduled Future (*Next Christmas falls on a Thursday*); Subordinate Future (*If I see him, I'll send him home*) etc. Each context was also coded for remoteness, following Dahl (2000a), with non-remote defined as 'today, tomorrow'. Finally, the grammatical person (1st, 2nd, 3rd person) associated with each verb form was also recorded.

Time Periods

The data was divided into 16 time periods, four per school year, instead of the two periods per school year used for the analyses of lexical development and accuracy of language use. It was decided that a more fine-grained analysis was necessary here to allow, for example, a more detailed description of the way existing form-meaning associations are affected by the introduction of new ones, as when emerging forms are overgeneralised, temporarily affecting the appropriate use of other forms.

Part Three ANALYSIS & DISCUSSION

6 Aspects of Language Development – Lexical Diversity and Accuracy of Language Use

6.1 Introduction

The main focus of this study is an analysis of the learners' acquisition of past and future temporal expression, which will be dealt with in Chapters 7 and 8. However, in a longitudinal study it is important to provide as detailed a picture as possible of the participants (Ellis and Barkhuizen, 2005) and Chapter 6 contributes towards this end, presenting some general background information on the individual learner's language development in relation to lexical diversity and overall accuracy of language use. Section 6.2 deals with lexical development – with the range and variety of vocabulary employed by the learners. Section 6.3. looks at overall accuracy and at the relative frequencies of different kinds of errors. It identifies areas where significant progress is made as well as those areas that remain problematic. Section 6.3 also contains an analysis of self-corrections and other changes made in the course of writing. This analysis provides some insights into the learners' orientation to the writing tasks which provided the written data – on the extent to which they were focused on meaning rather than on form. Points of particular interest are briefly discussed as they arise and the principal findings are summarized and discussed further in Chapter 9.

6.2 Lexical development

6.2.1 Introduction

Section 6.2.3 presents the results of a longitudinal analysis of the development of lexical diversity, an analysis conducted using the *vocd* program described in Section 5.4.1. First, however, Section 6.2.2 gives a general picture of the range and variety of vocabulary employed by the

learners over the course of the study. An indication of vocabulary size is provided by an analysis of the number of different words (types) used by each learner (Table 6.1) and then frequency patterns and patterns of development are briefly discussed in relation to two of the major word classes – verbs and adjectives.

6.2.2 Vocabulary

Luisa chose to write more often than the other two learners, and usually wrote at greater length; as a result she produced 25,379 words (tokens) in the analysed data, compared to 21,307 for Sara and 17,810 for Ana. Table 6.1 provides the types and tokens totals, and type/token ratios (TTR) for nouns, verbs, adjectives and adverbs. Of the major word classes, the learners display the greatest variety in their use of adjectives, and the least in their use of verbs. Ana, for example, only uses 70 different verbs in Period 1, and almost half of all verb uses in that period (47%, or 161/342) were accounted for by just four verbs – *copula be, go, want* and *like*. Her repertoire expands over time and at the end of the study, in Period 8, Ana uses 105 different verbs and the four most frequent verbs (copula *be, go, get, think*) account for 38% of all verb uses (153/398). In Period 8. Sara's four most frequent verbs (*copula be, think, have, go*) also account for 38% of all verb uses (253/669) and Luisa's top four verbs (*copula be, have, do, get)* for 34% of all verb uses (221/646).

Table 6.1 MAJOR WORD CLASSES - TYPES & TOKENS

	NOUNS			VERBS			ADJECTIVES			ADVERBS		
	types	*tokens*	TTR	types	*tokens*	TTR	types	*tokens*	TTR	types	*tokens*	TTR
LUISA	**1163**	*4685*	0.248	**498**	*4594*	0.108	**494**	*1655*	0.298	**158**	969	0.163
SARA	**729**	*4038*	0.181	**417**	*3888*	0.107	**353**	*1160*	0.304	**136**	678	0.201
ANA	**708**	*2986*	0.237	**282**	*3483*	0.081	**331**	*1124*	0.294	**101**	606	0.167

(proper nouns and auxiliary verbs excluded)

When auxiliaries are included in the analysis, the verb *be* (copula and auxiliary) accounts for more than a quarter of all verb uses by the three learners. *Be, have* and *do* together account for 40% of Luisa and Sara's verb use, and 38% of Ana's. These three verbs are also by far the most frequent in L1 English. In fact there is a striking similarity between the verb frequencies found in the learners' data and those found in the British National Corpus (Leech et al., 2001). Most of the learners' top 20 verbs are also found in the BNC's list of the 20 most frequent verbs – 17 of Luisa's, 16 of Sara's and 14 of Ana's, although their respective ranks vary (Table 6.2). And a number of the most frequent verbs on the BNC list appear just outside the learners' top 20 - *should* and *take* (Luisa) *should* and *would* (Sara) and *take, make* and *could* (Ana).

Table 6.2 Rank frequency list of verbs (by lemma)

rank	BNC	LUISA	SARA	ANA
1	be	be	be	be
2	have	have	have	have
3	do	do	do	do
4	will	would	go	go
5	say	go	think	think
6	would	get	will	say
7	can	could	see	will
8	get	can	say	get
9	make	make	get	come
10	go	think	tell	tell
11	see	look	can	want
12	know	see	make	could
13	take	come	come	like
14	could	will	want	can
15	think	give	take	see
16	come	know	know	start
17	give	find	could	ask
18	look	seem	like	look
19	may	want	ask	know
20	should	say	look	find
21	use	tell	feel	make
22	find	take	would	could
23	want	use	should	take
24	tell	show	talk	live
25	must	should	use	use

BNC 20 in red, BNC 21-25 in blue BNC data: Leech et al 2001

A comparison of verb frequencies at the beginning of the four year study (Period 1) and at the end (Period 8), shows that half or more of the individual learner's top ten verbs in each period are from the BNC top ten, and almost all are from the BNC top 20, with five verbs common to all three learners' 'top ten' in Period 1 (*be, have, do, go, come*) and six verbs common to all three 'top tens' in Period 8 (*be, have, do, get, make, think*). The main difference between the two periods is that auxiliary verbs are more frequent in the data in Period 8, with *can* for example, making it into Luisa and Sara's top ten, and *will* which is absent from Ana's data in Period 1, appearing as her 5th most frequent verb in Period 8.

In all, the 20 most frequent verbs in English account for 71% of Luisa and Sara's verb use, and 69% of Ana's. This is not really surprising given that these verbs are likely to be very frequent in the input which the learners received, facilitating early acquisition, (Ellis, 2012) and many frequent verbs are polysemous (Crossley et al., 2010) and therefore very useful for learners, enabling them to communicate a range of meanings – as the following examples of *know* from Sara's data illustrates:

[1] I didn't know how to do cycling. (Sara, file 12)

[2] One of her uncle already know one of the film star when he was in Gulf so they meet when they came here . (Sara, file 19)

[3] News channels will let us know about what's happening in the other side of the world or in another continent through satellite . (Sara, file 24)

[4] I know I am going to die or something worse is waiting for me in the darkness . (Sara, file 27)

The learners exhibit a much greater variation in their use of adjectives. Luisa's 20 most frequent adjectives account for only a quarter of all her adjective use – for Sara and Ana it is one third. And only nine adjectives are common to all three 'top 20' lists *(good, old, same, different, first, new, hard, big important*) six of these appearing among the BNC's 20 most frequent adjectives.

Ana's use of the adjective 'nice' provides an interesting snapshot of lexical development. It is her third most frequent adjective, whereas it does not feature in the BNC's top 50, and is rarely used by Luisa or Sara (typically accounting for less than 0.5% of their adjective use in any period). In Periods 1 and 2 Ana employs *nice* frequently to describe people, places and experiences, the word accounting for 11% of all adjective use.
For example, in a short retelling of the encounter between Cinderella and Fairy Godmother, Ana uses *nice* five times (see page v for list of codes):

[5] This woman was very *nice* and she had a magic wind [: wand] .... Her dress was very [/] white and very *nice* . Then Godmother gave for her glass slipper. It was very *nice*, so *nice* ... and then she changed pumkin [: pumpkin] on *nice* big coach. (Ana, file 13)

But from Period 3 on, *nice* accounts for 2% or less of all adjective use, as Ana's lexical repertoire broadens and she begins to employ a greater range of words to provide positive descriptions – adjectives such as *kind, pretty, beautiful, enjoyable, delicious* etc.

Luisa, Sara and Ana are exposed to different kinds of English language input – from conversations with their peers (both L1 English and L2 English speakers) on the one hand for example, to the language of the classroom and the curriculum on the other. These different influences are evident in the learners' use of lexis and particularly so in their use of adjectives, which range from the informal [6-9 below] to the decidedly formal [10-13]:

[6] I walked quite fast, I was *pissed off* and calling my cousins names. (Luisa, file 39)

[7] He's already *fed up* for playing referee betwee [/] during the fight between me and my mum . (Sara, file 68)

[8] This house was *fantastic*! (Ana, file 65)

[9] They are no longer people but robots who do exacly [: exactly] what they saw [//] see on TV and what TV says is *cool* . (Luisa, file 33)

[10] I moved on and changed my *adolescent* passion into a real thing . (Luisa file 77)

[11] There is also teenagers that can be very *idealistic* and caring, getting involved in voluntary community work (Ana, file 76)

[12] If they are sleep deprived they are usually *monosyllabic* during the next day . (Sara file 67)

[13] The writer used an *accesible* [: accessible] vocabulary. (Ana, file 83)

As might be expected from the research on prototypes and concept learning which suggests that basic category terms are acquired first (Rosch et al., 1976; Tomasello, 2003), it is generally in the later stages of the study that the less frequent, more formal or technical words tend to be found. In Periods 7 and 8 the following verbs , for example, are used for the first time in the data (and used appropriately):

Luisa: *symbolize, resemble, assure, oxygenate, compose, contrast, sympathise, exaggerate, practise, hydrate*

Sara: *motivate, convey, correct, direct, overreact, overtake, solve, interfere, interrupt, delete, provoke, approve, hesitate*

Ana: *prioritise, analyse, document, memorise, revise, impress, illustrate, migrate, judge, promote, rely, connect, campaign*

And among the nouns used for the first time are: *nightfall, ideologies, fantasy* (Luisa) *motherland, peer, breeze* (Sara) *gesture, curiosity, despair* (Ana).

The learners' vocabularies also continue to grow in depth as well as breadth, with polysemic words occurring with new meaning associations right up to the end of the study. In the last month of the fourth year for example, the following verbs are used in a new sense: *take, put* (Luisa) *make* (Sara) *know, get* (Ana).

6.2.3 Lexical diversity

To obtain a general picture of the learners' developing lexical proficiency, an analysis was conducted using the *vocd* program (see 5.4.1) , as research suggests that lexical diversity as measured by *vocd (D)* is the best predictor of lexical proficiency (Crossley et al., 2011) . 245 transcripts within the stable range of 100-400 tokens were analysed, 78 for Ana, 75

for Sara and 92 for Luisa. Only Luisa produced enough spontaneously written transcripts for a mean D analysis in Period 6.

The individual learner's D scores vary within each period. The variation is not simply – or necessarily - related to the type of writing task, but can involve more complex factors including for example, the number of topic changes in a transcript (Duran et al., 2004). However the mean D score for all the transcripts sampled in a period does provide a useful overall indication of lexical diversity.

Occasionally it is possible to make a direct comparison between two individual transcripts, as when Ana chose to return to two childhood memories she had previously recounted. Her second version of how she discovered there was no Santa has a D value of 55.47 compared to 36.66 for the first version written three months earlier. And the second version of her learning to ride a bicycle story has a D of 50.21 compared to 29.43 for the first, written over a year earlier.

The following excerpts from the bicycle stories give an indication of the difference in lexical proficiency between the first and second versions ([/] and [//] denote that the preceding word or scoped phrase was replaced by the subsequent word(s)):

[14] I always know that my father is [//] was behind me . I was riding until I saw that my father lets me go myself . When I saw that my father wasn't beside [//] behide [: behind] me then I felt [* fell]. (Ana, file 12, D = 29.43)

[15] First my dad was [/] always walked with me and was behind me . But once he let me go a [/] by myself. Then I didn't knew it but I [/] suddenly [/] I realised that he was one hundred metres <far than> [/] from me . It was a tragidit [: tragedy] ! I'd started to panic so I aga [/] felt [* fell] again.
(Ana, file 45, D = 50.21)

While there are few such opportunities for direct comparison in the data, the mean D scores for each period do provide clear evidence of development over time (Figure 6.1, Table 6.3).

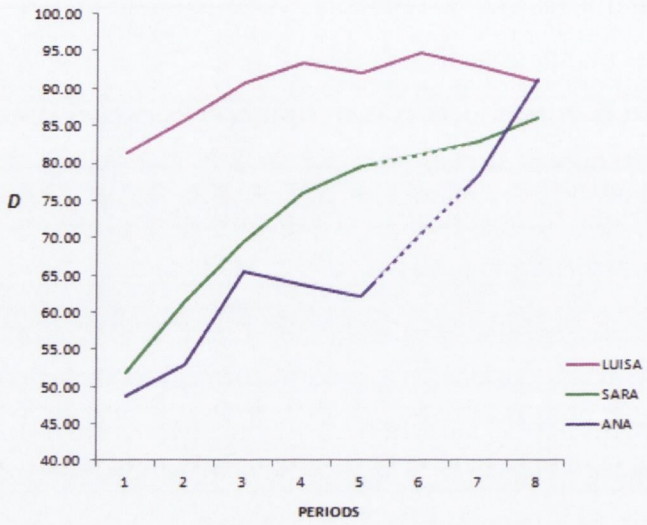


Figure 6.1 Lexical Diversity (*D*) by period

Table 6.3 Mean *D* values by period

	1	2	3	4	5	6	7	8
LUISA	81.43	85.72	90.82	93.35	92.00	94.79	92.95	90.99
SARA	51.73	61.36	69.59	76.08	79.48	**	82.83	86.04
ANA	48.76	52.89	65.39	63.67	62.19	**	78.53	91.09

** = insufficient transcripts for a mean D analysis in this period

A number of studies have included calculations of D values for oral recordings and written texts produced by L2 learners of English (e.g. Daller et al., 2013; Jarvis, 2002; Tavakoli and Foster , 2011; Yu, 2010). However most of these studies are evaluations or comparisons of different measures of lexical diversity or models of vocabulary learning, or are investigations of the effects of task design on learner production, and present the D data in ways which do not allow for a useful comparison with the data presented here. As a general indication however, a number of research findings may be noted. Duran et al. (2004) report a mean D of

56.58 for transcripts obtained from recordings of 32 pre-intermediate level adult learners of English engaged in an oral decision-making task. Yu (2010) reports a mean D of 73.00 for compositions produced by 200 candidates for the MELAB English language exam. The candidates were adult L2 English speakers taking the exam for the purposes of professional certification or college admission. In the case of 25 candidates, Yu also analysed transcripts of face-to-face interviews and found little difference between the mean D scores for written and spoken output (76.29 and 74.11 respectively).

Information on trends in the development of lexical diversity for L1 written English is reported in Malvern et al. (2004). Narrative compositions by 311 seven year olds, 383 eleven year olds, and 224 fourteen year olds were analysed using the *vocd* program. The mean D values obtained were 39.48 (age 7), 65.81 (age 11) and 73.05 (age 14). The 14 year old students who were rated most proficient writers overall, with scores of 6 or 7 on a scale of 1-8, had a mean D of 80.60 and 84.81 respectively. Duran et al. (2004) report a mean D of 90.59 for adult L1 writers, based on an analysis of samples of academic writing.

As Table 6.3 shows, Luisa obtains high mean D values right from the beginning of the study. In Periods 1 and 2 her lexical diversity is comparable to that of the 14 year old proficient writers in Malvern et al's L1 English study. From Period 3 on (the first part of the second year), the lexical diversity of her writing is comparable to that of adult L1 writers. In Period 1 Sara and Ana have mean D values lower than those found for the pre-intermediate level L2 learners in Duran et al. and lower than the D of 53.59 reported for the most proficient 7 year old L1 writers in Malvern et al. Sara's lexical diversity increases significantly over the first two years of the study and she makes steady progress after that, obtaining a mean D score of 86.04 in Period 8. Ana's D development is different, with a significant rise in Period 3 followed by several periods of stagnation, the most progress being made towards the end of the study. Her *minimum* D

values rise steadily in each period however, from 25.87 in Period 1 to 61.11 in Period 8. In Period 8 Ana achieves a mean D of 91.09 comparable to that of L1 adult writers.

Luisa's D for Period 1 seems very high, given that she had only been learning English for one year at this stage. It could be due, at least in part, to the fact that she was a proficient writer in her L1 (see Section 4.6.2). In the year before the study commenced , and subsequently, Luisa was attending regular Polish classes and, as a result, was continuing to develop her literacy skills in her L1. Cummins (2000) has proposed the existence of a Common Underlying Proficiency (CUP) for languages and, related to this, a Linguistic Interdependence Hypothesis which suggests that literacy skills developed in the L1 will transfer to the L2. Luisa may have been able to transfer to English the writing skills she acquired in Polish, deploying her learner English vocabulary in a way which minimized repetition (and by minimizing repetition increased her lexical diversity score).

Luisa's data contains a higher number of less frequent word types than is found in the data of the other two learners. 160 of the verbs she uses, for example, are below 500 on the BNC frequency list, compared to 114 for Sara and 61 for Ana. It is possible that her continuing L1 vocabulary development supported the development of her English vocabulary. Some of the less common, more formal or technical words which appear in her written data are very similar to their Polish lexical equivalents and may have been acquired first in Polish, facilitating their acquisition in English (eg: *alpinist/alpinista, ideology/ideologia, symbolize/symbolizować, transform/transformacji, terrorize/terroryzowac* etc).

From the age of 11 in Sara's case, and 12 in Ana's, their literacy development had been taking place almost entirely through their L2 . It is possible that Luisa's strong – and still developing- L1 literacy skills gave her an advantage over the other two learners when it came to the

development of lexical diversity in English. However, as the *vocd* analysis shows, Sara and Ana do continue to make progress in this area over the four years of the study, ultimately attaining (or in Sara's case approaching) the adult L1 English level.

6.3 Accuracy of language use

6.3.1 Introduction

On its own, an accuracy analysis would be a very negative way of looking at a learner's interlanguage, because a lot of 'errors' are inevitable, even necessary features of developmental processes, and can be indicative of progress (Devitt, 1992). However an analysis of the learners' grammatical and lexical ability to perform accurately in the L2 can nevertheless contribute to the overall picture of language development (Ellis and Barkhuizen, 2005). This section presents the results of an *errors per 100 words* analysis which was used to chart the development of the learners' overall accuracy of language use (see section 5.4.2) , and supplementary information is provided on the relative frequencies of different kinds of errors, and on the areas – like the article system – which remain problematic. The section begins however with a discussion of the learners' self-corrections and changes, specifically with the degree to which the learners, while engaged in these writing tasks, were primarily concerned with the communication or clarification of meaning (a focus on meaning) rather than on the achievement of accuracy (a focus on form).

6.3.2 Focus on meaning versus focus on form

As already outlined in sections 4.4.3 and 4.5.4, this study aims to sample implicit linguistic knowledge rather than explicit knowledge, to the extent that that is possible with written tasks. With this goal in mind, the data was elicited through a process designed to encourage relatively spontaneous language use. The learners were encouraged to focus on

meaning, and on communicating as much as possible in a relatively short period of time, typically 10-15 minutes in the initial stages and then, as the learners capacity for sustained writing increased, 20-25 minutes. This researcher, as the learners' language teacher, was always present when these texts were produced and formed the general impression that the learners wrote quickly, only infrequently pausing to review or reflect on their writing. The changes which the learners made in the course of writing were examined to obtain an indication of the extent to which there was a focus on meaning.

Because they are unplanned and written quickly, the learners' texts often contain false starts or immediate rephrasings (coded [/], see p.v) resulting, for example, from a change in the meaning they wish to convey or an elaboration or clarification of the meaning [examples 16-18, final phrasing is highlighted here in bold]. Sometimes a learner decides to make a change when they are in the middle of the writing of a word, as is the case with the second change in example 17.

[16] < I have a house in Vilnius> [/] <I have> [/] **In Lithuania I have a home in Vilnius** . (Ana, file 11)

[17] They [/] **Some of them are blown from branches** <are on the ground or s > [/] **and are under** the [/] **trees or are** flying [/] **pushed far away by wind** . (Luisa, file 12)

[18] **Then the godmother saw the cat who was Ella's best friend and told to go and get** some [/] **six rats** . (Sara, file 13)

And there are immediate corrections to grammar [19-21] and also, but very rarely, to lexis [22] :

[19] They both loves [/] love each other but afraid that they will lose their head . (Sara, file 41)

[20] My brother is sweet but I don't think I'm going to have childrens [/] children in the future . (Luisa file 4b)

[21] Then I sat on this bicycle and ride [/] I was riding very good. (Ana, file 12)

[22] A man in un [/] a blue uniform ordered [//] orders us to put the gougles [/] goggles on . (Luisa, file 25)

Occasionally one error is replaced by another [23-24]:

[23] I flay [/] flyed plain [: plane] Aer Lingus. (Luisa, file 4b).

[24] It was that a [/] the old woman who I met in the forest. (Ana file 61)

And sometimes an error is introduced in the course of an unnecessary correction:.

[25] When old year is [/] are changing at New Year . (Ana file 10)

About one fifth of the corrections (coded [//]) were made after the completion of the phrase or sentence, but in the course of the initial writing of the text (as explained in Section 4.5.4 the task of carefully 'reading over' the text and making corrections was done at a later stage, usually in the next class, and only the initial drafts are included in the data). These corrections typically involve the insertion of omitted articles [26] or changes to the verb [27-28] .

[26] When I was small and still belived [: believed] in Santa I was writing letters to him week [//] a week before Christmas. (Luisa, file 11)

[27] Does Juliana <likes fly> [//] like to fly ? (Ana, file 16)

[28] The snake is [//] was heavy when he caught the snake and tried to carry it . (Sara, file 7)

The analysis reveals that a very large majority of the learners' changes - between 79% and 90% - involve a change in the meaning or in the way the meaning is communicated. As Table 6.4 shows, attempts to correct errors of grammar, syntax or lexis, or to correct forms wrongly assumed to be errors, only account for 10% overall of Luisa's changes , 11% of Sara's and 21% of Ana's.

Table 6.4 Changes and corrections by period

	LUISA		SARA		ANA	
	[/] per 100 clauses	% err corr	[/] per 100 clauses	% err corr	[/] per 100 clauses	% err corr
PERIOD						
1	54	20%	14	10%	31	24%
2	46	11%	21	15%	34	36%
3	55	9%	19	19%	44	16%
4	47	7%	18	13%	53	17%
5	54	10%	18	11%	39	27%
*6**	34	10%	36	3%	40	21%
7	27	9%	27	8%	31	17%
8	33	5%	37	8%	41	9%
mean	*44*	*10%*	*24*	*11%*	*39*	*21%*

[/]per 100 clauses = number of changes ([/]+[//]) per 100 clauses

% err corr = % of changes ([/]+[//]) involving error correction

6 small samples in this period for Ana and Sara (<100 clauses)*

As Ellis and Barkhuizen (2005) observe, the opportunities learners have for self correction depends on the number of errors they make – the more errors, the more opportunities. A low percentage of error-focused changes could just reflect a low error count. However this is not the case here. In fact the vast majority of errors go uncorrected at the time of writing , on average 96% of Sara's, 90% of Luisa's and 89% of Ana's (Table 6.5).

Table 6. 5 Errors not corrected by period

	1	2	3	4	5	6	7	8	*mean*
LUISA	84%	90%	89%	89%	87%	90%	94%	93%	*90%*
SARA	98%	95%	97%	96%	96%	97%	97%	95%	*96%*
ANA	94%	88%	91%	87%	84%	85%	90%	96%	*89%*

When they carefully read over and corrected their writing a day or a few days later, all three learners were almost always able to identify and correct at least some of the errors they had made, and occasionally a majority of them (typically agreement errors, omitted Simple Past *–ed* etc). While it is probable that some errors were avoided through the learners employing their explicit knowledge in the course of writing, overall this analysis of changes and corrections does seem to indicate that they were more focused on meaning than on accuracy when producing the samples used in the data, and this is important where the aim is to sample, as far as possible, implicit rather than explicit knowledge (see Section 4.4.3).

6.3.3 Development of accurate use

Global Measure of Accuracy

In terms of overall accuracy, there is a considerable gap between Ana and the other two learners at the start of the study (Figure 6.1, Table 6.6). Ana's *error per 100 words* rate (16.7) is more than double that of Sara (7.3) or Luisa (7.5) in Period 1 and it is only in Period 4, in the second half of the second school year, that her rate is comparable to those obtained for the other two learners in Period 1. After this Ana continues to make progress, but more slowly, with several periods where little change occurs.

Table 6. 6 errors per 100 words by period

	1	2	3	4	5	6	7	8
LUISA	7.5	5.3	4.6	3.6	4.1	3.5	3.1	2.4
SARA	7.3	6.4	6.5	5.4	5.6	5.6	4.3	5.0
ANA	16.7	13.9	9.8	7.2	7.2	7.3	6.0	6.4

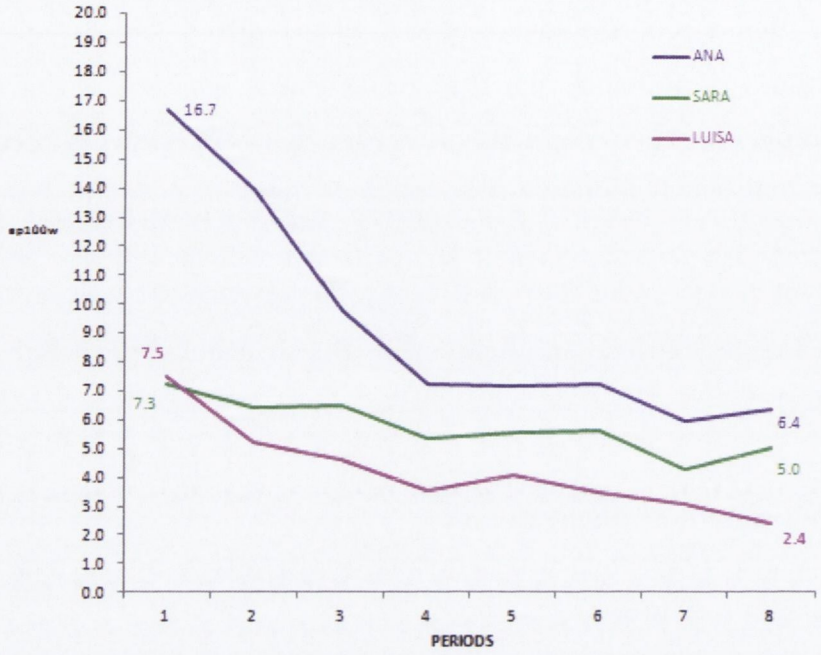


Figure 6.2 Errors per 100 words, by period

By the end of the four years Luisa has achieved a very low error rate of 2.4. The biggest reduction in her error rate occurs in Period 2. Arguably, over time, the achievement of big reductions becomes less likely as the higher the level of accuracy achieved, the more difficult it becomes to improve on it, both because of continuing development in the verb system, and because what remains to be acquired elsewhere is increasingly likely to relate to the more problematic features of grammar and lexis – like certain aspects of the article system or preposition use. Although her overall accuracy of language use is still higher than Ana's in Period 8, Sara makes relatively little progress over the course of the study, with some errors – in areas such as nominal and verbal agreement - persisting to the end.

Roughly a quarter of Ana and Luisa's errors are found in the verb system, and over a third of Sara's. Most occur in the course of the development of the learners' tense-aspect-mood system which is discussed in Chapters 7 and 8.

Verbal Agreement

Agreement remains a problem area for Sara throughout the study and accounts for 26% of all her verb-related errors, compared to 18% for Ana and 16% for Luisa. In Period 8, at the end of the study, more than a third of Sara's verb errors involve agreement (21/57). At this stage Luisa and Ana are only very occasionally making mistakes with *be* and the 3rd person singular present.

[29] This is a wonderful description about things we do and we have in our life, which *is* the ingredients of life .
(Sara, file 76)

[30] Two advices that I would give to a thirteen year old is one, accept the advices that your parents *gives* you, two have an aim in your life . (Sara file 70)

[31] Yes, this poem *make* me feel hopeful, the hope for the future, the day they're waiting for their dreams come true .
(Sara file 82)

Negation

Negatives account for 8% of the coded clauses in the data (6% of Sara's, 8% of Luisa's and Ana's). As Cancino et al. (1978), Wode (1976) and others established, negation develops gradually in stages towards the norms of the target language (see section 3.3 for a discussion) From the start of the study Luisa's uses of the negation system is over 90% accurate. However Sara and Ana initially have difficulty with auxiliary do forms. Sara uses tense and person markers on *do* from Period 1 [32, 33] suggesting that she had already passed through the unanalyzed *don't* stage of negation development:

[32] Romeo is lovesick because Romeo loves Rosaline but she *doesn't love* back . (Sara file 8)

[33] Luisa *didn't* go [/] *went* to summer school . (Sara, file 4a)

However she only rarely uses do-negatives and only very rarely uses them accurately, until the end of the second year, when she suddenly starts to make extensive use of the form, getting it right 95% of the time. Ana marks *do* for person but not for past tense in Period 1 [35]. Some of the forms she uses in this period [34, 36] might suggest that she was still at a relatively early stage, but as acquisition research has shown, learners do not leave one developmental stage behind when they enter another; at a given point in time learners may use forms typical of several different stages.

[34] They *not live* in the sea because there is too much salt . (Ana, file 7)

[35] But he *doesn't saw* this letter . (Ana, file 9)

[36] Bird *doesn't to want* to leave him. (Ana, file 6)

During the first two years of the study Ana attempts do-negatives much more frequently than Sara and there is a steady increase in accuracy, but she does not achieve 95% accurate use until Period 6, the end of the third year.

Questions

Unless they are specifically elicited, questions are likely to be much less frequent in written data than in recordings of spoken English. Luisa's written samples contain 75 direct questions, Sara's 126 and Ana's 99. There were a few writing tasks which called for questions (an interview for example, or a learner-devised revision test for an exam text) but narratives [37] and opinion pieces [38] also produced contexts for questions.

[37] Was I imagining things that I wanted to happen ? (Sara, file 50)

[38] Isn't it better that your children can mess up when they still can expect your help ? (Luisa, file 84)

From Period 2 on Luisa and Sara are producing questions with the auxiliary in second position after a wh-word [39-40] which is generally considered indicative of the later stages in the development of the interrogative in English Pienneman, Johnston and Brindley, 1988).

[39] Why does it have to be like that ? (Luisa, file 14)

[40] When did she leart [: learn] swimming? (Sara, file 15)

Ana does not begin to do so until the middle of Period 3 but once she does her use is 95% accurate – she is, in fact the first to master do-questions :

[41] Why didn't I start to study for my Mocks earlier? (Ana, file 40)

Luisa's use of the form is over 90% accurate from Period 4, towards the end of the second year. At this stage Sara generally gets the syntax right, but in Period 4 although all her negative *do* statements are accurately formed, about half her *do* questions have both the auxiliary and the verb marked for tense [42-43] something which Luisa and Ana did occasionally at the start.

[42] Did she had pain from collarbone ? (Sara, file 15)

[43] What did your friends said when you told them this idea ? (Sara, file 30)

It is only in the fourth year of the study that Sara routinely uses an accurate do-question form.

Articles

The English article system can be problematical for learners, especially for those whose L1 does not have articles (Luk and Shirai, 2009) – as is the case for Ana, Luisa and Sara. While frequency is an important factor in acquisition, so also are salience and the form's reliability as a predictor of an interpretation (Ellis and Collins, 2009), and though the articles are among the most frequently used forms in English, they have low salience and there is no one-form-meaning correspondence to aid the learner. The choice of *the, a* or no article 'depends on a complex interaction of factors including meaning, shared knowledge, context and whether the noun is singular, plural or uncountable' (Parrott 2010, 25)

Overall article errors accounts for 27% of Luisa's errors, 25% of Ana's and 19% of Sara's. In Period 1, after a year in an English language

environment, Ana's article use is only accurate 36% of the time. She makes steady if slow progress and by the fourth year of the study, her accuracy has increased to 79%. Luisa's accuracy rises from 55% in Period 1 to 81% in Period 2 and after that fluctuates between 79% and 92%. Sara begins with a relatively high rate of accuracy, at 79% which then fluctuates between 76% and 92%. For all three learners article use is most accurate in the final year of the study.

In Period 8 Ana is still making relatively simple errors, omitting articles (oa, othe) before singular nouns, often, but not always, when there is an adjective before the noun:

[44] It is really *oa* simple thing if you get oan education, you will get *oa* better job (Ana file 79)

[45] I think the princess feels sad that she felt in love with *oa* servant, not a prince. (Ana file 85)

[46] It gives me *oa* clear picture of what is being described (Ana file 81)

Most of Ana's article errors in Period 8 (19/26) are of this type . Most of Sara and Luisa's however involve more complex uses:

[47] Another reason is , it is based on *the* personal experience so the writer know what he is talking about. (Luisa file 91)

[48] It seems like *the* nature is getting ready for another ending of the day and welcoming the night. (Sara, file 77)

[49] Never let anyone become *othe* centre of your life when you are only an option to them . (Sara, file 70)

[50] There are only *oa* few weeks left before your Leaving_Cert exams are starting. (Luisa, file 96)

Once an accuracy rate of around 80% has been achieved, further progress seems to be slow, possibly because – as suggested by Luisa and Sara's data - it is the more complex uses which remain to be acquired.

Prepositions

Prepositions are also problematical for learners. They often have to make choices and distinctions that are not necessary in their L1 and many of these choices or decisions may appear to them to have very little or

nothing to do with meaning. Prepositions account for 11% of Ana's errors overall, 13% of Sara's and 12% of Luisa's, and are still presenting a similar range of difficulties at the end of the study in Period 8, as the following examples with *on, in, into* and *to* illustrate:

[51] I started to chat with my friends in [/] *in* Facebook and Bebo . (Sara file 70)

[52] They can be a form of relaxation that is substitute *to* watching television (Luisa, file 99)

[53] If God put a hope *to* our hearts it means it's achievable. (Ana file 83)

[54] If that too failed he was imaning [* imagining] himself fall *into* his belly and drag himself with his elbows until his strength was gone . (Sara, file 74)

[55] I think it's [/] the teenage years is like preparation *to* 'big' life . (Ana file 79)

[56] It was too late to [/] to get *on* [* to] my interview . (Luisa file 88)

[57] I used to like to [//] going there and feeling the wind blowing *in* my cheeks. (Ana, file 65)

Lexical errors

Lexical errors (wrong words, malformed words etc) are relatively infrequent in the data, comprising 5% of Sara's errors, 7% of Ana's and 11% of Luisa's. The errors have a wide range of sources and each specific error is generally only made by one of the learners, usually just once [58-61]. However all three learners occasionally confuse *go* and *come* , an error relating to perspective and position, and common enough to feature in grammar books [62,63].

[58] The animal doctor will examinate [* examine] the pigeon. (Luisa, file 7)

[59] I'm worry [* worried]. (Ana, file 28)

[60] Tennis always last [* ends] in a quarrel . (Sara, file 10)

[61] It is the time when you are between the infancy and the adulcy [* adulthood] so teenagers years are natural to make mistakes. (Ana file 75)

[62] The young prince decided not to come [*] back to Corinth so he could change his faith [* fate]. (Luisa file 67)

[63] When we were sailing back suddenly my friend with two boys went [*] over and splashed us with water . (Ana file 35).

6.4 Conclusion

The analyses presented in this chapter provide clear evidence of the learners' English language development, in relation to lexical development and overall accuracy of language use, continuing for the four years covered by the study. Sara and Ana began with low lexical diversity scores – lower than the mean D recorded for proficient 7 year old L1 writers, but by the end of the study Ana has achieved, and Sara is close to achieving , a level of lexical diversity found in adult L1 writing. Luisa achieves this level in the middle of the second year, however there is evidence of continuing development in her English vocabulary right up to the end, with new words and new form-meaning associations appearing in the final month. Sara's and Ana's data also show evidence of their vocabulary continuing to expand, in depth as well as breadth, over the full four years.

All three learners achieve improvements in their general accuracy of language use over the course of the study, however there are individual differences in the rate of progress and in the final outcomes. Ana, who was at an earlier stage of language development than the other two, started with a high *error per 100 words* rate of 16.7, but ended with a rate of 6.4 – a marked improvement, although it was largely achieved in the first two years after which further development was slow. Luisa and Sara started with similar error rates (7.5 and 7.3 respectively) but while Luisa ended the study with a very low rate of 2.4 errors per 100 words, Sara made less progress, and her rate (5.0) was double that of Luisa's in the final period.

The data also shows the learners successfully progressing through a number of developmental sequences with Luisa usually completing the process first, followed by Sara - though this is not always the case. Ana for example is the first to achieve over 90% accuracy in the use of do-questions.

It seems to require a long time for some aspects of the system to be acquired. The article system remains a main source of errors for all three learners, with Ana still making relatively simple article errors in Period 8. And all three continue to experience some difficulty sorting out the form-meaning associations of prepositions. Agreement is an area which is persistently problematic for Sara. In Period 8 she makes 41 errors in relation to nominal and verbal agreement. In comparison Luisa makes 5 such errors in that period and Ana 11.

In summary , all three learners make considerable progress in the area of lexical proficiency, with developments continuing up to the end of the study; and while individual differences are evident, general accuracy of language use continues to increase over the course of the study with all three learners achieving their most accurate use in the final year.

7 The Acquisition of Past-Related Temporal Expression

7.1 Introduction

Chapter 6 presented some general background information on the learners' language development and discussed in particular the extent to which the learners' lexical diversity and general accuracy of language use increase over time. However the main focus of this study is an analysis of the learners' acquisition of temporal expression, specifically tense-aspect morphology, and the discussion of that analysis begins here. This chapter deals with past-related expression and Chapter 8 will look at future expression.

The expression of temporality plays an important role in all linguistic communication and has been one of the main areas of research in first and second language development (see Section 3.5 for a discussion). The Simple Past is the first past form to emerge in the interlanguage of learners of English (Klein et al., 1995) and in her studies of the development of the Past Progressive and the Present Perfect, Bardovi-Harlig used the development of the Simple Past as a benchmark or ' system-specific means of gauging development and comparing the learners' (2000, 142). In this chapter, the learners' use of the Simple Past is examined first (7.2) both to provide an important part of the picture of the learners' overall language development, and as a prerequisite for the discussion of later-emerging forms. Section 7.3 looks at the Past Progressive and Section 7.4 at the Present Perfect.

In the case of all three forms the analysis begins by looking at the rates of appropriate and accurate use in required contexts, before going on to consider the ways in which the forms are employed - in particular the spread of these forms from prototypical to less typical uses. Some points of particular interest are briefly discussed as they arise and the principal findings are summarized and discussed further in Chapter 9.

7.2 The Simple Past

7.2.1 Past time contexts and the Simple Past

Past is the most common time context in the data (Figure 7.1). The writing tasks from which the data is drawn were not designed to elicit any particular forms or contexts. Instead they sampled the range of tasks which the learners were asked to undertake in the mainstream English classroom. Therefore it is possible that the breakdown here is representative of the time contexts which would naturally arise in the course of an L1 student's written assignments.

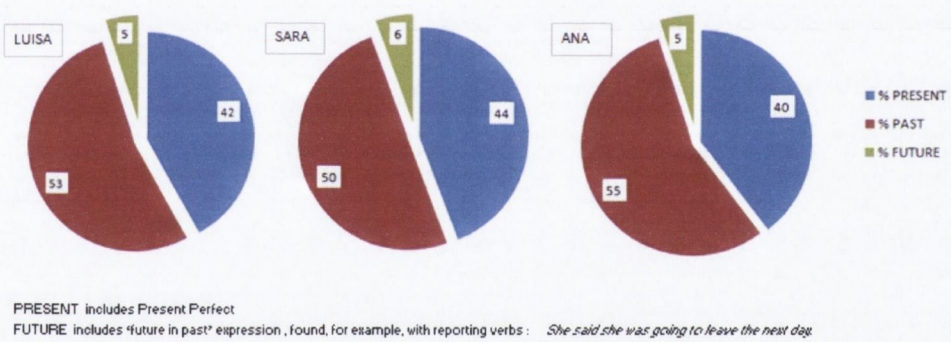


Figure 7.1 Past, Present and Future contexts in the data

The Simple Past is by far the most common of the past contexts . Including interrogatives and negatives (where tense marking is generally carried on the auxiliary rather than the main verb) it accounts for 89% of Luisa's past contexts, 90% of Sara's and 91% of Ana's. If interrogative and negatives are excluded, the Simple Past still accounts for 89% of all past contexts. This is in line with corpus findings that simple verb phrases (present and past) account for about 90% of all verb phrases (Biber et al.,1999).

7.2.2 The development of appropriate & accurate use of Simple Past

Not surprisingly given its high frequency in the input which learners receive, and its general utility, the Simple Past has been found to be the first past form to emerge in the acquisition of English as a second language (Bardovi-Harlig 2000). Because of the particular research questions which guide this study, it was decided to choose learners who had reached the stage in the development of tense-aspect morphology where they were close to acquiring the Simple Past (see Section 4.6.1).

At the start of the study Ana and Luisa had been living in an English language environment for one year and Sara for two years. Given that 80-90% appropriate use is often considered to indicate acquisition (Ellis and Barkhuizen, 2005) and that Luisa and Sara's rates of appropriate use are almost always above 80% (Figure 7.1, Table 7.1), it would appear that the Simple Past was already well established in the tense-aspect systems of these learners.

From Period 3 (in the second half of the first year of the study) Ana's appropriate use rate is nearly always well above 80%. As the ESF study (Klein et al., 1995) observed , the development of morphology is slow and gradual, with appropriate and inappropriate usage co-existing for a long time before full control of the functions of a form is achieved and Ana's usage rates continue to fluctuate into the fourth year as, for example, new forms emerge in her tense-aspect system. To a lesser extent Sara's rates also fluctuate during the first two years.

Table 7.1 Use of Simple Past periods 1-10 (of 16)

	LUISA		SARA		ANA	
	%APP	%ACC	%APP	%ACC	%APP	%ACC
PERIOD						
10	*100%	*100%	100%	100%	100%	100%
9	96%	90%	94%	94%	88%	84%
8	95%	95%	84%	82%	*79%*	*77%*
7	97%	93%	*76%*	*76%*	90%	90%
6	94%	88%	*100%	*100%	96%	92%
5	94%	89%	81%	81%	88%	88%
4	90%	83%	90%	90%	80%	80%
3	90%	84%	87%	87%	88%	80%
2	97%	91%	*74%*	*74%*	*64%*	*61%*
1	87%	83%	89%	*79%*	*75%*	*66%*

%APP = % appropriate use (includes overregularised forms, eg: telled, thinked)

%ACC = % accurate use *= small sample (less than 15 different verb types)

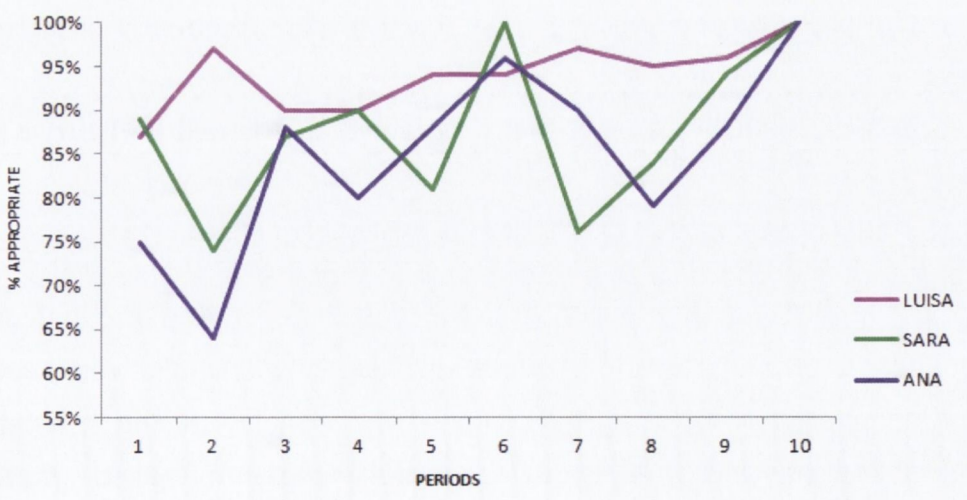


Figure 7.2 % Appropriate use of Simple Past periods 1-10

Luisa

1436 Simple Past contexts (involving 227 different verbs) were analysed in the case of Luisa, whose appropriate usage was the most stable, never dropping below 87% (and from Period 7 on, always over 95% appropriate and accurate). In the early stages – the first half of the first year of the study – her use of the Simple Past is not perhaps as stable as the appropriate use rate suggests, as there are quite a number of instances where she only 'gets it right' after initially supplying an inappropriate form. In the following examples [/] indicates an immediate correction and [//] a later correction (see p. X for list of codes):

[64] She just *go* [/] went to bed . (Luisa, file 4a)

[65] The Scylla *hiden* [/] hid in the cave (Luisa, file 6)

[66] All he *want* [//] wanted to do wa [/] was leave [//] to leave the pigeon and to play football again . (Luisa, file 7)

[67] At first we *sing* [//] sang Christmas carols . (Luisa, file 11)

There is no obvious change over time in the kinds of forms which occur in the instances where Luisa fails to supply an accurate Simple Past when required. From Period 1 on, base verbs are used and there is also overgeneralization of the regular *–ed* to irregular verbs. Base forms are used with both common verbs [68] and also with infrequently occurring ones [69]:

[68] One night at twelve our uncle *come* and moved [/] he tryie [/] tried to scared us . (Luisa, file 4a)

[69] Scylla *shriek* and swooped . (Luisa, file 6).

Apart from *maked,* employed three times in Period 1, and one instance of *gived,* overregularisation involves low-frequency irregular verbs each used just once or twice in Luisa's data - it is possible that the input she received contained few tokens of these verbs in the Simple Past. *fly* for example only occurs 90 times per million words in the British National Corpus (BNC) and it is possible that up to the point where she goes to use it in a Simple Past context the input Luisa received had not included

many examples of *flew*. She guesses an irregular form, before settling on an overregularised one:

[70] I *flay [/] flyed* plain [: plane] Aer_Lingus. (Luisa, file 4b).

The other verbs which are overregularised are all less frequently occurring (in the BNC) than *fly* – for example *wake, hide, strike, shoot, swear* [72] and *freeze* [71].

[71] The light went on. The children *freezed*. They saw a shadow in one [/] the window facing backyard . (Luisa, file 26)

[72] This is when I *swared* to myself ' I'm gonna [/] going to design and build bikes someday . (Luisa, file 77)

So at this stage of the development of Luisa's Simple Past, most errors involve infrequently occuring verbs, or the occasional overgeneralization of emerging perfect forms (to be discussed later in Section 7.4).

Sara

Sara's data provided 1165 Simple Past contexts for analysis, involving 166 different verbs. Her rate of appropriate use is generally high –though it falls to 74% in Period 2 and drops again in periods 5 and 7. She occasionally hesitates over an appropriate form [73] or provides an appropriate form only after first using an inappropriate one ([74] but in general she is less likely to self-correct in the course of writing than either Luisa or Ana (see Table 6.4)

[73] I think it was [/] was a great adventure in my life . (Sara, file 19)

[74] The snake is [//] was heavy when he caught the snake and tried to carry it . (Sara, file 7)

When she fails to supply an accurate or appropriate form in a Simple Past context Sara almost always uses a base/present form of the verb. She rarely overregularises – in fact she does so only twice at the very beginning of the study, both times with the verb *sleep*:

[75] She *sleeped* in a tent near a forest . (Sara, file 4a)

[76] She *sleeped* at home rest of the days . (Sara, file 4b)

Most of Sara's inappropriate use involves verbs which are frequent in the language – *come, have, take, make, think, walk, happen, stop* etc – and which she uses appropriately and accurately in the Simple Past both before, as well as after, the incidences of inappropriate use:

[77] When he saw the special pigeon Cherokee he went to take her and he slowly *take* her out of the nest. (Sara, file 5)

[78] As I started to the new way I *think* I heard footsteps at the back . I told myself ' don't <be panick> [//] panic . As I *walk* my childhood memories began to work . A story my grandmother told me, a same situation like this, a girl got killed in the darkness. (Sara, file 27)

Only a few instances of inappropriate usage involve less frequently occurring verbs like *hate* and *notice*:

[79] Jem said not to annoy their aunt anymore but Scout *hate* it when Jem told her what to do . (Sara, file 18)

[80] One dancer fall down from [/] on the stage while she was dancing but noone saw it or *notice* it because she made <a step> [//] her own step . (Sara, file 19)

The newly emerging present and past perfect forms are overgeneralised to Simple Past contexts, but this happens only very occasionally (and will be discussed in Section 7.4). From Period 10 on, Sara's use of the Simple Past is always 95% or above, appropriate and accurate. Its use is occasionally overgeneralised to perfect contexts from Period 12 on (and this too will be discussed in Section 7.4)

Ana

1149 Simple Past contexts, involving 146 different verbs, were analysed in Ana's data. With an appropriate use rate of 75% in Period 1, dropping to 64% in Period 2, she seems to be at an earlier stage of development than the other two learners are at the start of the study. Both accuracy and appropriacy rates rise to above 80% after that, until Period 8, the end of the second year, when there is a noticeable drop coinciding with the emergence of the Present Perfect and its overgeneralization to Simple Past contexts. There is another decrease in appropriacy and accuracy in

Period 13, at the start of the fourth year of the study, partly due to the overgeneralization of perfect forms.

In the first year of the study, two forms each account for around 50% of inappropriate use in Simple Past contexts. The first is the base/present :

[81] When Romeo knew that, he *run* and killed Tybalt. (Ana, file 9)

The second is *was* + verb –ed (and occasionally *was* + base). In a couple of instances this form is found in ambiguous contexts, where the Past Progressive might also have been appropriate [82] but generally it is found in Simple Past contexts, occurring with state, activity and event verbs [83] .

[82] She and her friends went to Funtansia . She [/] There she *was played.* (Ana file 4a)

[83] I remember my New Year when I was a small girl. I b [/] *was believe* that Santa Close [/] . I *was waited* for prize ..... and she *was puted* all toys in red bag and when was New Year we got the same this prizes. (Ana, file 10) .

This form was used by Sara and Luisa occasionally in the year before the study commenced, when their Simple Past development was at an earlier stage. It appears only once in each learner's data in both cases with the verb *keep*:

[84] I tried a lot but I *was keep* falling from the cycle . (Sara, file 12)

[85] She *was keep* saying how great there was. (Luisa, file 13)

Ana's pattern of inappropriate use changes from Period 4 on, with her now most likely to overgeneralise other past-related forms, initially the Past Progressive [86] and then perfect forms – particularly in Period 8 where the emerging Present Perfect is used 11 times in all, 8 times in a Simple Past context [87]. Where she uses a base/present form it is slightly more likely to be a very common verb (*have go run* etc) than an infrequently occurring one (*ruin, remind*).

[86] I caught the spider then I put to my little pocket and I *was saying* for my mum 'that is for your gift'. (Ana, file 17)

[87] ' What to do?' he asked me. *I've told* him 'you can go outside out'. 'Can I?' he asked again. *I've looked* out of window. I knew that I *haven't suppose* to let him go by himself but then 'what would happen with him?' I thought. (Ana, file 49)

From the second year of the study on (from periods 5 to 16) Ana's appropriate use rate is 88% or higher, apart from periods 8 and 13 where overgeneralisation of perfect forms depresses the rate to 79%.

7.2.3 Lexical aspect & the learners' use of the Simple Past

The Aspect Hypothesis (Andersen and Shirai, 1994) suggests that the first verbs to attract Simple Past marking will be event verbs (achievement verbs like *arrive* and accomplishment verbs like *build*) whose inherent meaning implies an end point. This hypothesis has been substantially supported by research. Both Bardovi Harlig and Reynolds (1995) and Collins (2002) for example, found that in cloze tests learners showed higher use of past with event verbs than with activity or state verbs. Collins also found that learners continued to be influenced by lexical aspect even as their use of past tense morphology became more productive.

Bardovi-Harlig (2000) notes that the spread of the Simple Past to state verbs has been difficult to observe in free written production (as opposed to cloze tests) because learner production is often quite limited, with *be* and *have* the most frequent statives. However there is sufficient data in this study to establish that Ana, Sara and Luisa had all reached a stage of morphological development where both activity and state verbs are routinely found with past marking.

In fact, in the first half of year one of the study when Ana's appropriate use rate was 75% initially, falling to 64% (Table 7.1), 20% of the verbs found with past tense marking were state verbs [88] and 34% were activity verbs [89]. This suggests that the spread of the form to activity and state verbs does not depend on the stability of the Simple Past.

[88] We *liked* this school but one thing what I [//] we don't like that was a lot of homework . (Ana, file 4b)

[89] In the river she *swam* and *played* . (Ana, file 4)

Even though the data shows that Simple Past marking has spread to all four verb classes, there is still some evidence of the influence of lexical aspect continuing over the first two years of the study. On the infrequent occasions when a learner supplies a Simple Present/base verb or *–ing* participle in a Simple Past context, in 77% of cases (47/61) it is a state or activity verb.

Over the four years of the study, while only about 10% of the verbs used by the learners are state verbs, they are among the verbs most frequently found with Simple Past marking. As Table 7.2 shows, seven of Luisa and Sara's 'Top 20' are state verbs, and nine of Ana's are.

Table 7.2 Rank Frequency of Simple Past Verbs

rank	LUISA	SARA	ANA
1	*be*	*be*	*be*
2	*have*	go	go
3	go	*have*	*have*
4	look	*see*	come
5	say	say	say
6	get	tell	ask
7	come	come	start
8	find	*think**	*think**
9	make	ask	tell
10	*think**	take	*see*
11	*want*	make	get
12	*know*	get	*want*
13	*seem*	*feel*	decide
14	decide	*want*	meet
15	*see*	decide	look
16	tell	do	take
17	ask	find	*know*
18	start	*know*	find
19	spend	happen	*hear***
20	take	pass	*like***
			*realise***

state verbs in *blue* * = some uses of *think* are not stative

** these verbs are ranked joint 19th

Most of these state verbs are in fact among the most frequently used verbs in English (Leech et al., 2001). As Wulff et al. (2009) discovered, an analysis of L1 corpora reveals that the same common high-utility verbs dominate the frequency list in each tense-aspect category (see Table 3.1).

7.2.4 High-frequency, high-utility verbs

As already mentioned, one of the criteria used to select learners for this study was that they had reached a stage of morphological development where the Simple Past was being used appropriately at least 70-80% of the time. While the learners' early use of Simple Past markings may have been heavily influenced by lexical aspect as research suggests (Bardovi Harlig and Reynolds, 1995; Collins, 2002), from the start of this study they are using verbs from all lexical aspect classes in Simple Past contexts, most of the time with appropriate and accurate marking. Overall the frequencies in the learners' Simple Past output reflect the L1 frequencies in the BNC. Of the ten most frequent verbs in the Simple Past category in the BNC (*be do have say get go think come take and want*) seven are found in the top 12 verbs of each learner (Table 7.2)

Boyd and Goldberg (2009) say that learners tend to hear the same set of high-frequency, high utility words over and over again. This is a good characterization of the three learners' Simple Past output. Of the 227 different verbs used by Luisa, the ten most frequently occurring account for more than half of all contexts (53%), while at the other end of the scale, 99 verbs (or 43% of all verbs) are used only once. Sara's pattern of use is very similar. She employed 166 verbs in all, and her ten most frequent account for 54% of all contexts, while 41% of verbs are used only once. Ana, whose Simple Past development was the least advanced at the start of the study, relies a little more on her 'Top Ten' verbs – they account for 61% of all contexts – but the percentage of verbs used only once, 44%, is very similar to that found for Luisa and Sara. For all three learners

verb use in Simple Past contexts follows a Zipfian distribution (see Section 3.5 for a discussion).

7.2.5 The Spread of Simple Past from prototypical to less typical uses

Andersen and Shirai (1996) hypothesized a sequence for the internal structure of the category past tense, from prototype to marginal members. In it, habitual and iterative uses follow use with state verbs, and are in turn followed by counterfactual use. They suggested that the sequence was unlikely to be strictly linear but rather hierarchical, with habitual use for example beginning to develop slowly once past marking had spread from achievement to accomplishment verbs.

Simple Past with habitual meaning [90, 91, 92] is found in all three learners' data from the beginning of the study, but, as Andersen and Shirai suggest might be the case, it is a developing use and does not become firmly established for some time - until the start of the second year for Luisa and Ana, and late into the second year for Sara. Up to this point they occasionally use the base form [93] or the progressive [94] in contexts where an L1 speaker would have preferred the Simple Past.

[90] She always *asked* advices and she has followed my advices. (Sara, file 56)

[91] My mother always *told* that I have to be good because [: because] Santa just comes to the good children . (Ana, file 14)

[92] I spent summer holidays in Ireland looking after my little brother. I woke [/] was [/] *woke up* every day at seven thirty and I *worked* as babysitter . (Luisa, file 4b)

[93] Whenever he *do* that a shock of terror would run through my spine. (Sara, file 35)

[94] When my little legs were tired, every time I sat on this pig and she *was going* to home. (Ana, file 21)

During the first two years the learners also employ *used to* or, occasionally, *would +verb* in contexts where the Simple Past could have been used with habitual meaning [95, 96]. Sara is the most likely to do so and this use of alternative (but also appropriate) forms may partly explain why the habitual Simple Past takes longer to become established in her interlanguage.

[95] When I was a small girl, about four years old, every summer I *used to go* to Belarus to visit my grandmother. (Ana, file 21)

[96] Sometimes Angel *would smile* at me. (Sara, file 35)

According to Andersen and Shirai the least typical use of the Simple Past form is the counterfactual use , in conditional constructions such as [97]. While it is not a past time use, and therefore outside the scope of this particular discussion, it is nonetheless interesting to note that Luisa and Sara use it appropriately from around the time the habitual Simple Past has become established in their interlanguages (in line with Andersen and Shirai's prediction). Ana only creates 4 contexts and overgeneralises the *would* + verb construction in each [98].

[97] If I *ended up* on a desert island and I could pick the character I'd like to spend time with , I'd pick Rita. (Luisa, file 49)

[98] If I *would be* stuck on desert she would <know how> [/] cook [/] make delicious food . (Ana, file 42)

7.2.6 Summary

The Simple Past is well established in Luisa and Sara's data from the start of the study and from Period 3, in the second half of the first year, in Ana's data. Luisa's use of the form is the most stable; from Period 2 it is always over 90% appropriate, and from Period 7 always over 95% appropriate. As Klein et al. (1995) observed, the development of morphology is slow and gradual, and Ana's appropriate usage rates continue to fluctuate

into the fourth year of the study, and Sara's until the end of the second year. The learners routinely use Simple Past marking with all four lexical aspect classes. However there is some evidence of the lingering influence of lexical aspect during the first two years of the study, with inappropriate use most likely to occur with activity and state verbs. A small number of high-frequency, high-utility verbs dominate Simple Past contexts and in general the learners' use of verbs follows a Zipfian distribution. As predicted by Andersen and Shirai (1994) the habitual use of the Simple Past appears early on, but develops slowly. It does not become firmly established in the learners' interlanguage until the second year of the study.

7.3 The Past Progressive

7.3.1 Introduction

The progressive is most typically used to designate an event or state of affairs in progress or continuing (Biber et al., 1999). According to Leech (2004) it is increasing in frequency in English, especially in the spoken language , but is still much less frequent than the Simple Present and Simple Past. The progressive occurs in between 5% and 10% of verb phrases, depending on register, compared to 90% for the simple tenses (Biber et al., 1999) . In the learners' data the progressive accounts for 6% of past-related contexts – a similar frequency profile to that found in L1 use.

In her longitudinal study of the emergence of past-related morphology Bardovi Harlig (2000) found that the Simple Past and the Past Progressive developed at around the same time in her learners' interlanguage and she suggests that the emergence of the Past Progressive is dependent, not on the stability of the Simple Past, but rather on the prior emergence of the Present Progressive.

7.3.2 The Development of appropriate and accurate use

The Present Progressive was clearly well established in the learners' tense-aspect systems before the study commenced. There are 78 present progressive contexts in Luisa's data, and 94 in Sara's, and the form is used appropriately in all. Ana creates 57 contexts for the Present Progressive and only twice fails to supply a Present Progressive form.
Initially the Past Progressive seems to be underused in Sara and Ana's data as Luisa creates 15 contexts in Period 1, but Sara and Ana only create 2 each. However in the data as a whole all three learners demonstrate a high rate of appropriate use in Past Progressive contexts – 95% for Luisa, 92% for Sara and 94% for Ana - and their pattern of appropriate and accurate use suggests that the form was probably well on the way to becoming established before the study commenced.

In all Luisa creates 84 Past Progressive contexts . There are four instances of inappropriate forms being used, none of these instances occurring in the early stages of the study. Two involve the overuse of the Present Progressive, but two are more complex – involving less prototypical (habitual) contexts for progressive (and will be discussed later in this section). There are 77 Past Progressive contexts in Sara's data, with 7 instances of inappropriate use spread across the four years of the study, involving either the use of the bare progressive or the overuse of the Present Progressive. Ana creates 85 Past Progressive contexts, failing to supply the appropriate form in five of them, using the perfect participle with auxiliary *be* on two occasions and Simple Past twice (once in a habitual context).

7.3.3 Verb frequencies in Past Progressive

While a small number of high-utility, frequently occurring verbs dominate the learners' use of the Simple Past in the data, this is not the case with the Past Progressive. Luisa uses 58 different verbs in her 84 Past Progressive contexts, and Sara uses 40 in her 77 contexts, and in both

cases only three verbs occur more than three times, with Sara's most frequent verb (*walk*) occurring seven times and Luisa's (*sit*) occurring five times . In her 85 Past Progressive contexts Ana uses 43 different verbs, with only four verbs occurring more than three times, with the most frequent verb (*walk*) occurring eight times. Only two verbs are used more than three times by all three learners – *sit* and *go*.

There are a relatively small number of Past Progressive contexts in the data (246 tokens between the three learners), probably too small for a meaningful analysis of frequency patterns in past contexts. However it can be noted that, of the six verbs which occur more than three times in a learner's data, four (*walk, sit, look* and *wait*) are among the ten verbs found by Wulff et al. (2009) to be most distinctively associated – as opposed to most frequently associated - with progressive marking in the BNC (see Table 3.2). An analysis of all verb uses with progressive marking (1,213 tokens) reveals that the frequencies in the learners' output reflect L1 frequencies in the BNC. In the case of each learner, seven of her ten most frequent verbs are also among the ten most frequent in the BNC progressive list , with six being common to all three learners – *go, do, be, get, come, try*.

7.3.4 Lexical Aspect and the learners use of the Past Progressive form

The Aspect Hypothesis (Andersen and Shirai, 1994) predicts that the progressive will be found initially with activity verbs, eventually spreading to event (achievement and accomplishment) verbs. A number of researchers - e.g. Bardovi Harlig (1998), Collins (2002), Housen (2002) and Huang (1999), have found evidence of a distributional bias towards activity verbs in the interlanguage of L2 learners of English, and Housen reported that in the earlier stages of acquisition the use of the *–ing* marker was mainly restricted to activity verbs.

Luisa, Sara and Ana are in the later stages of the acquisition of the Past Progressive and while activity verbs account for a majority of contexts, it is only a small majority. Activity verbs [99-101] are found in 55% of Luisa's Past Progressive contexts , 52% of Sara's and 57% of Ana's (which is similar to the percentages found for present progressives in their data – 54%, 53% and 53% respectively). In comparison, Huang (1999) found that in her L1 oral data, 46% of all progressive-marked predicates occurred with activity verbs.

[99] The thought was torturing me, me and my friends going to die . (Sara, file 32)

[100] Maria was speaking to me but I was thinking about this woman , couldn't stop. (Ana, file 61)

[101] He said they were gambling and he won Jem's trousers . (Luisa, file 26)

Almost all the rest of the verbs found in the learners' Past Progressive contexts are event verbs [102-3]. Although some state verbs can combine with the progressive aspect where an 'activity' meaning can be supplied (Leech 2004) there are only two such instances in the data [104-5].

[102] I went to buy a ball while she was changing her cloths [: clothes]. (Ana, file 56)

[103] I was telling her I would never pass my driving licence [//] test. (Luisa, file 87)

[104] All way to that park I was remembring [: remembering] my precious time with her and all the musics she played for me . (Sara, file 56)

[105] In the second stanza he was imagining that the light between the ricks of hay and straw is [//] was like light from heaven . (Ana, file 62)

7.3.5 The spread of the progressive to repeated actions and events

In terms of the spread of the form from activity to event verbs it seems that the learners' use of the progressive is quite developed, approaching the L1 pattern of use. However the development of the progressive can also be examined in terms of the extension of its use from prototypical to

less typical meanings. According to Andersen and Shirai (1996) this would be from process to iterative and then to habitual or futurate meanings and finally to stative progressives.

In her analysis of the form's occurrence in two L1 corpora, the BNC and the Bank of English, Romer (2005) found that in English the progressive is used much more frequently with an iterative/habitual meaning than had previously been thought. She reports that the progressive is used around 39% of the time to report repeated, as opposed to single, events or actions. And she suggests that 'repeatedness' should move further to the centre of descriptions of the English progressive.

Luisa, Sara and Ana all employ the progressive for repeated situations from early on in the study, but overall they do so relatively rarely. 'Repeatedness' [106-8] features in only 10% of Luisa's use of the Past Progressive, 9% of Ana's and 4% of Sara's (Sara's use is higher in Present Progressive contexts at 10%, Ana's and Luisa's about the same for both).

[106] Every minute I *was calling* my brother but t [/] I couldn't find him an [/] nowhere . (Ana, file 52b)

[107] When I was small and still belived [: believed] in Santa I *was writing* letters to him week [//] a week before Christmas . (Luisa, file 11)

[108] Last week our newspaper reported an incident happen to a club, that someone keeps [/] keep breaking glasses of the window . They placed cctv cameras around the club building . Finally they found who *was doing* this . (Sara, file 29)

The low rate of use of the progressive with repeated meaning is not due to a low incidence of 'repeated action or event' reports in the data, but rather to the learners creating contexts for other forms or constructions. The learners may use the Simple Past [109] or *would* + verb [110] or '*used to*' [111] or other means [112] to convey the iterative or habitual nature of the situation.

Sometimes, as in [109, 110, 112] the Past Progressive (with an adverbial) could have been used instead.

[109] My mum always *said* that about teenagers and she wasn't wrong saying they are irresponsible. (Ana file 77)

[110] They would always take me along with them which ment [: meant] I had to move at least twice a year (Luisa, file 65)

[111] I used to get lots of presents for my [/] the Christmas because I am the youngest girl in the whole family . (Sara, file 9)

[112] I'd started to panic so I aga [/] felt [fell] again . It wasn't first time and last . (Ana, file 45)

There is one instance in the data where a learner can be observed in the act of choosing between the Past Progressive and another form in a habitual context. Here Luisa initially used the Simple Past but subsequently rephrased :

[113] He <always made her mad> [//] *was always driving her mad* by his crazy ideas of fun (Luisa, file 50)

7.3.6 Summary

So while the three learners usually supply the Past Progressive where the context they have created requires it, with an appropriate use rate of over 90%, and while the form is used almost equally with activity and event verbs, the learners are still in the process of acquiring the less typical, but still fairly common, 'repeatedness' function of the progressive, with its particular meanings such as 'habit in existence over a limited period' and 'repetition of events of limited duration" (Leech, 2004). Because they have other means of expressing 'repeatedness' it is possible that the learners need to be exposed to a considerable amount of L1 input in order to fully acquire the L1 speaker's contrastive use of the 'repeatedness' meanings of the progressive and other forms.

7.4 The Present Perfect

7.4.1 Introduction

With the emergence of the Present Perfect, learners are able to use morphological means to link past and present time, and so, in a sense the appearance of this form represents an important milestone on the developmental path of temporal expression. According to Leech (2004) there are two distinct ways in which a past event may be related to the present by means of the Present Perfect - it may involve a time period leading up to the present and it may have results persisting at the present time.

Leech distinguishes four different uses of the form:

1. a state extending over a period leading up to the present moment: *I've lived in this neighbourhood since I was a kid.*
2. 'indefinite past' use referring to an event which has occurred 'at least once in a period leading up to the present' : *I've known love, but not true love.*
3. habitual or iterative use , referring to 'a habit in a period leading up to the present': *I've always walked to work.*
4. 'resultative past' where the form is used in reference to a past event to imply that the result of that event still pertains, or as Leech puts it, is still operative at the present time: *Someone has broken her doll* (the doll is now broken).

7.4.2 The emergence of the Present Perfect

According to Bardovi-Harlig (2000) the Present Perfect is the third past-related form to emerge, with its emergence depending on the stability of the Simple Past in the learner's interlanguage. In her analyses, a tense-aspect form is not considered to have emerged until it has been used appropriately with three different verbs . On this basis, the Present Perfect

emerges relatively late for the learners in this study - the end of the first year for Luisa, early on in the second year for Sara and at the end of the second year for Ana.

In the case of Luisa and Sara, the Present Perfect makes its first appearance in the data at the end of Period 2, around the middle of the first year of the study. In Period 2 Luisa has a rate of appropriate use of the Simple Past of 97%, and Sara of 74% (however in both cases Period 2 was not typical - Sara's rate is 89% in Period 1 and 87% in Period 3; Luisa's is 87% in Period 1 and 90% in Period 3). In Bardovi Harlig's 1997 study, the rate of appropriate use of the Simple Past ranges from 68% to 93% at the time of first recorded use of the Present Perfect, with a mean group rate of 86% for all 16 learners.

The Present Perfect does not appear in Ana's data until Period 7, in the middle of the second year of the study. At the time her Simple Past appropriate use rate is 90%, with an average rate of 88% for the previous year. In Ana's case, the first appearance of the Present Perfect does not follow soon upon the achievement of high rates of appropriate use of the Simple Past. And in the case of all three learners, the emergence of the Present Perfect, its use with three different verbs, occurs after the Simple Past has been stable for close to, or over, one year.

Even after its emergence, the Present Perfect is used infrequently by Luisa and Sara, and even less frequently by Ana. However generally when they do create a context where the Present Perfect would be preferred by an L1 writer, the learners tend to supply an appropriate form in that context (an appropriate form being defined as auxiliary *have* with a main verb, but not necessarily the perfect participle) and also usually an accurate one. The learners seem to have little difficulty with the form of the Present Perfect in fact. Counting all uses of the form, including overgeneralisations in non-Present Perfect contexts, only 7% (6/86) are

malformed, (which is the same percentage found in Bardovi Harlig's written data).

Luisa creates 34 contexts, and supplies an appropriate form 27 times, an appropriate usage rate of 80%. Sara creates 33 contexts and supplies an appropriate form 29 times (87%) and Ana uses the form appropriately 11 times in 14 contexts (79%). The Simple Past is by far the most common alternative form supplied in Present Perfect contexts possibly because of the degree of semantic overlap between these two forms (which will be discussed later in this section). A competing awareness of the present and past features is seen in [114] where Luisa initially goes to use a Simple Present form in a negative Present Perfect context, before opting for the Simple Past form. And the difficulties the learners experience in sorting out when present relevance requires the Present Perfect, and when it does not, can be seen in [115] where Ana is discussing an opinion she now holds on the basis of an interview she has just read.

[114] I heard he was quite good actor but for some mesterious [: mysterious] reason he *dose [/] didn't do* it since I remember. (Luisa, file 27)

[115] When I was reading this interview *I've noticed* that he is poetic and sensative [: sensitive] person by looking at things he writes about . (Ana, file 69)

7.4.3 The learners' use of the Present Perfect

According to Leech (2004) , of the four principal meanings of the Present Perfect, 'resultative past' is by far the most common, followed by 'indefinite past' with the remaining two ('state up to the present' and 'habit up to the present') being considerably less frequent. The learners employ the form to express all four meanings and the relative frequencies found in L1 usage are also reflected in the learners' data. Resultative past [116-118] accounts for 56% of all Present Perfect use and indefinite past [119, 120] for 35%.

[116] A lot of people *have learned* to speak a foreign language by following courses on television. (Ana, file 75)

[117] I think princess [//] the princess is leading an ordinary life but *something's made* her sad. (Sara, file 40)

[118] The motorbike driver *have just crashed* his dreams. (Luisa, file 30)

[119] I have heard older people saying three things in life that won't come again, time, opportunity and words. (Sara, file 22)

[120] I've never seen a thing like that in real life, only on tv, but it's not the same thing, right? (Luisa, file 69)

The remaining two, less typical, meanings are only used by Luisa and Sara and they appear late - from Period 10 in the middle of the third year of the study in Luisa's case, and from Period 15 in the fourth year, in Sara's:

[121] Have you always been interested [: interested] in mechanics [: mechanics] ? (Luisa, file 77)

[122] If you haven't done any work at all until now, nine weeks of hard work can make the difference to your results. (Sara, file 78).

Ana only creates one context for habitual use at the very end of the study, and uses the Simple Present instead of the Present Perfect.

The learners' use of adverbials with the Present Perfect also reflects the L1 patterns of use reported by Leech (2004). Leech notes that with 'indefinite past' uses, the indefinite meaning is often reinforced adverbially, especially by *ever, never* or *before*. In 70% of the learners' 'indefinite past' contexts (16/23) an adverb is used – almost invariably either *ever* or *never* [123] but also *before* [124]. Adverbials of duration are usually required in 'state' and 'habit' contexts [125] , and the learners supply them , 4 out of 5 times with the context supplying the duration of the period in the 5th. The 'resultative past' meaning needs no support from adverbials, Leech says, and generally the learners do not use them – there is just one use each of both *already* and *just* by Luisa and Ana [126, 127].

[123] She was the (most) beautiful girl I *have ever met*. (Ana, file 67)

[124] I *have never passed* through a situation like this *before*. (Sara, file 50)

[125] I haven't seen her *for half a year* (Luisa, file 68)

[126] When we look closer we can see the tail of the other parrot who's *already* flown out of the window into the sunshine. (Luisa, file 64).

[127] The german officer who <just flew> [//] *have just flew* to Casablanca , was [//] is he really important officer and what [/] why did he get [//] fly to Casablanca? (Ana, file 58)

7.4.4 Overgeneralization of the Present Perfect

All three learners overgeneralize the Present Perfect to other tense-aspect contexts – in fact Ana's early use of the form is characterized by overgeneralization. The Present Perfect first appears in her data at the end of Period 7 and in Period 8 she uses the form eleven times – but only three times in a Present Perfect context. In total Ana overgeneralizes 11 times, Luisa 5 times and Sara 3. Of the 19 overgeneralisations, 16 occur where an L1 speaker would have preferred the Simple Past:

[128] <Did you> [/] *Have you ever been* afraid during that trip ?
(Ana, file 44, talking about a trip that had taken place some months before)

[129] She *have emailed* me his photo but I didn't look at all and pretended I didn't get the email . (Sara, file 50)

[130] *Have you done* mechaniks [: mechanics] in school ?
(Luisa, file 77, question addressed to a middle-aged man)

In [131] Ana initially uses the Simple Past, but then changes to the Present Perfect, possibly because she has related the past event to the present situation (with *even though*) and is paying particular attention to the 'present relevance' feature as the Present Perfect begins to establish itself in her tense-aspect system.

[131] I think the boy has a good relationship with his parents because [/] even though his father <didn't let> [//] *haven't let* him go on a hunting party to the Lagoons when he was under sixteen years old .

Something similar may be happening in [130] above, as Luisa is asking the man how he has ended up as a famous engineer.

Bardovi Harlig (1997) also found that her learners were more likely to overgeneralize the Present Perfect to Simple Past, rather than Simple Present or Past Perfect contexts. She suggests that this may be because the degree of semantic overlap is greatest for these two forms, that the two forms uniquely share truth value, the Present Perfect sentence being true in all cases where the Simple Past one is, and vice versa.

7.4.5 Summary

Research suggests that the Present Perfect emerges after the Simple Past and Past Progressive (see 7.4.2) and for these three learners it emerges much later, after the two past forms have been used appropriately and accurately for a year or more. Morphosyntactic complexity and semantic complexity have both been proposed as possible explanations for this later emergence (Bardovi-Harlig, 2000). The learners in this study have been routinely combining an auxiliary with a participle since the start, (the *be +-ing* of progressive constructions) and there is little evidence in this data of them struggling with the *have +ed/en* construction. The Present Perfect form is accurately produced 93% of the time (7.4.2). However the instances of non-suppliance, hesitation (changes made in the course of writing) and overgeneralization do provide evidence of the learners' struggling with the semantics of the Present Perfect (7.4.3). Semantic complexity may be an important factor in determining the order of emergence, especially when combined with frequency of input. The form is much less frequent than the Simple Past in L1 English (Biber et al., 1999) and, as a result, is probably relatively infrequent in the input the learners are exposed to.

Even after emergence, the learners use the Present Perfect relatively infrequently. Between them they create only 81 contexts and occasionally they do not supply the form where they have created the context - such instances occur up to the end of the four years of the study. In fact Sara and Ana both fail to supply the form in the very last Present Perfect

contexts they create, and Luisa does so in three of her final five contexts. It would seem, then, that after four years in an English speaking environment (five in Sara's case) the learners are still in the process of acquiring the form-meaning associations of the Present Perfect.

On the other hand, Luisa and Sara have started to employ the less typical habitual meanings of the form by the end of the study and Ana's use becomes more appropriately targeted, with less overuse in the later stages. Also, all three learners show an L1-like pattern of adverbial use with the form.

7.5 Conclusion

The Simple Past is well established in Luisa and Sara's interlanguage from the start of the study and in Ana's interlanguage from the middle of the first year. From the beginning, verbs from all four lexical aspect classes are routinely found with Simple Past marking. In fact Ana's data suggests that the spread of the form to state verbs does not depend on the stability of the Simple Past. The form appears with its less typical 'habitual' meaning early on, but as Andersen and Shirai (1994) predict, this use develops slowly. It does not become established until the second year of the study.

The Past Progressive is used with a very high degree of appropriateness and accuracy right from the beginning. While activity verbs are found in a majority of Past Progressive contexts, it is only a small majority – in this respect the learners' pattern of use approaches that of the three L1 speakers in Huang's 1999 study. However the data suggests that the learners have not fully acquired the less typical but still fairly common 'repeatedness' meaning of the Past Progressive – it accounts for 10% or less of their use of the form, in comparison to the 39% found for L1 speakers (Romer, 2005). As predicted by Andersen & Shirai (1996) stative progressives are very rare – in fact there are only two occurrences in the data.

The Present Perfect emerges at the end of the first year for Luisa, early on in the second year for Sara and at the end of the second year for Ana. In the case of all three, it emerges at a stage where they have been using the two past forms appropriately and accurately for more than a year. The learners' data suggests that semantic complexity may help to account for this late emergence. Even after emergence the learners create relatively few contexts for the Present Perfect, and late into the final year still fail to supply the form in some of the contexts they do create. By the end of the study, however, Luisa and Sara's use of the form has extended to its less typical 'state' and 'habitual' meanings.

The principal findings presented in this chapter are summarized and discussed further in Chapter 9.

8 The Acquisition of Future Expression

8.1 Introduction

Chapter 7 focused on the learners' past-related temporal expression, looking at the development of the Simple Past and Past Progressive, and at the emergence and development of the Present Perfect. Chapter 8 looks at the learners' future time expression. 8.2 begins by identifying the different environments in the data where the learners have the opportunity to create future contexts, and then provides an overview of the forms used in those contexts, the different ways in which they are employed and their relative frequency of occurrence (Figures 8.1 and 8.2) The developing use of each form is then examined, in the order in which it emerges in future contexts in the learners' interlanguage – the Will Future (8.3) the Subordinate Future Simple Present (8.4), the Futurate Present Progressive (8.5) and the Go-Future (8.6). Several minor future forms or uses are also examined, in each case in the section dealing with the most relevant major form – *will* +progressive infinitive in 8.3, Futurate Simple Present in 8.4 and the *be about to* construction in 8.6. The dominance of the Will-Future and the late emergence of the Go-Future are examined in 8.7. Other points of particular interest are discussed as they arise and the principal findings are summarized and discussed further in Chapter 9.

8.2 Future contexts & future forms

The writing tasks from which the data is drawn were not designed to elicit any particular forms or contexts (see 4.5.4 for a discussion) . Instead they generally sampled the range of tasks which the learners were asked to undertake in the mainstream classroom (the learners sometimes set their own tasks and topics, an option always available to them). Of the 267 texts included in the data, less than a dozen have a main theme relating to the future, and there are only three texts where future time expression predominates. However past-tense texts and present-tense ones do

provide opportunities for the learners to make reference to the future. Such opportunities arise, for example, when the learners are expressing opinions or giving advice:

[132] More importantly you *will* actually *find* yourself talking to your child instead of shouting (Luisa, file 84)

[133] We can hope that someone *will stand up* to save our planet (Sara, file 80)

[134] If *you'll take* our advices seriously.... *You'll do* much better and surely *you'll* achieve your goals. (Ana, file 84)

Responding to a literary text may involve speculating about a likely outcome [135] or a character's future behaviour [136] or feelings about the future [137]:

[135] I think she *will win* some day. (Ana, file 83)

[136] He *won't leave* him forever. He will come back to visit Dazza. (Luisa, file 7)

[137] Even though in these difficult time they do [/] are afraid that they *will lose* each other and their family might split (Sara, file 63)

And in personal and fictional narrative texts, reference to the future may arise in direct speech [138-9] reported speech [140-1] and in 'background' comment on a situation or event [142-4].

[138] 'I shall prepare you for a ball (Luisa, file 14)

[139] 'Well there must be somebody who's going to like me (Ana, file 66)

[140] The Oracle told the messenger that the baby will kill King Laius (Sara, file 53)

[141] My parents were furious at me and told me that I would have punishment for running from home . (Ana, file 64)

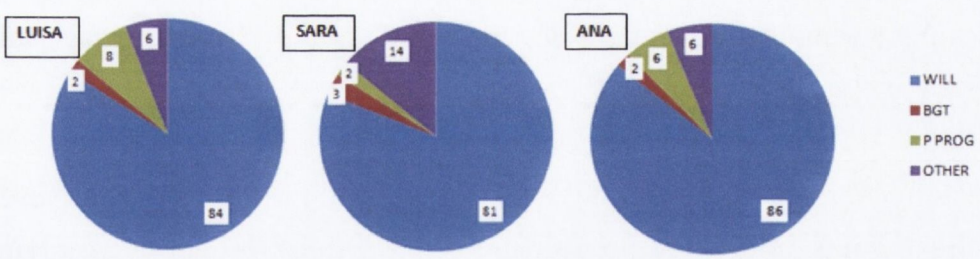
[142] This might be the last trip with my friends because we are going to graduate from school . (Sara, file 42)

[143] I'm not going to forget this place (Ana, file 65)

[144] Because they were coming next month when [//] and by this time everything would be unpacked. (Luisa, file 75)

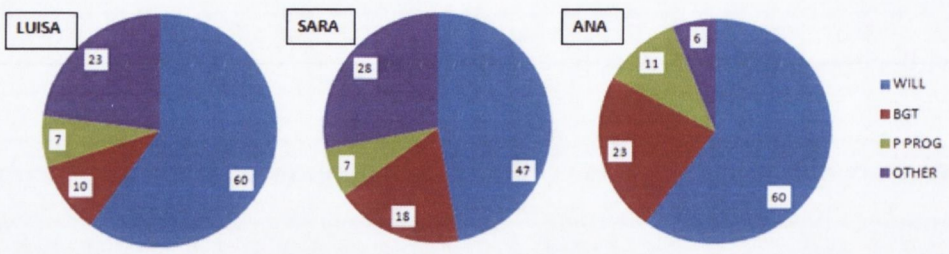
In practice only 5% of the contexts created by the learners are for future forms, including 'backshifted' future and future-in-the-past reference (see Fig. 7.1). However although this is a very small percentage, it still provides 468 future contexts for analysis.

As already discussed (3.6.2) English, in common with many other languages, has a variety of forms used in future expression. In this data the learners create nine different types of future context, from the 'neutral future of prediction' to the 'matter of course future', and use six different forms. By far the most common form is *will* +infinitive. Until the middle of the study the Will-Future accounts for more than 80% of all future contexts (Fig. 8.1). Even after the emergence of *be going to* (Fig. 8.2) the *will* form is still dominant, involved in 60% of Luisa and Ana's future time expression, and almost half of Sara's. The frequency order in Figure 8.2 is in line with that reported for L1 English by Leech (3.6.2) with the exception of Ana's infrequent use of the Subordinate Future Simple Present.



Will = *will* +infinitive BGT = *be going to* (Go-Future) P Prog = Futurate Present Progressive
'Other' includes the Subordinate Future Simple Present as well as some minor, infrequently occurring forms
emergence = use of *be going to* with 3 different verbs

Figure 8.1 Forms used in Future contexts before the emergence of the Go-Future % breakdown



Will = *will* +infinitive BGT = *be going to* (Go-Future) P Prog = Futurate Present Progressive
'Other' includes the Subordinate Future Simple Present as well as some minor, infrequently occurring forms
emergence = use of *be going to* with 3 different verbs

Figure 8.2 Forms used in Future contexts after the emergence of the Go-Future % Breakdown

In the data, prediction-based future time references outnumber intention-based ones by a ratio of 2.7:1, and 3rd person subjects are the most common (62%) followed by 1st person (24%) and 2nd person (14%). This is probably a reflection of the wide range of writing tasks undertaken by the learners. If the data were drawn mainly from learner journals or from recordings of conversations, it is likely that personal plans and intentions would feature more (as is the case in Bardovi-Harlig's study, for example, where the language samples came mainly from journals and guided conversational interviews).

The following sections examine the different future forms as they are used in the learners' interlanguage, starting with the most frequently occurring form, the Will-Future.

8.3 The Will Future

8.3.1 Introduction

Of the two main grammaticalized future forms in English, *will* evolved in the language first, acquiring its characteristic meaning of prediction and providing English with its closest approximation to a neutral or colourless future (Leech, 2004). Although *be going to* is becoming increasingly popular in informal English, *will* continues to dominate – it is the most frequently found form in future contexts. L2 acquisition research has found that *will* emerges first in the learners' interlanguage (Bardovi-Harlig, 2004a) and this is the case here. *will* + infinitive is used by Luisa and Sara from Period 1 and by Ana from Period 4. For the first two years covered by the study it is the form most frequently used for both prediction-based and intention-based future reference.

8.3.2 Emergence and developing use

Luisa

Luisa appears to have acquired the Will-Future before the start of the study. In the data she creates 114 potential contexts for the Will-Future and until Period 8, half way through the study, she always supplies an accurate *will* form. Then she begins to occasionally use '*would*' inappropriately in place of *will* [145-6]. She does so 10 times in all, the final three instances in Period 16 at the end of the study. This use emerges at a time when she has already been using modal *would* and backshift *would* appropriately for more than a year and there's no obvious pattern or trigger, except that in 6 instances it occurs with temporal, *so that* or *if* constructions [147] with which she was still occasionally having difficulty.

[145] The parrots are flying into an unknown world that's promising promising] them a better future. We don't know what *would happen* next but those birds are full of hope . (Luisa, file 64)

[146] Lydia has dubs [: doubts] about meeting Jonathan who used to be her sister's boyfriend . It's a_bit hard for her . She's sneaking into the house like a rubber [: robber] so her sis [/] Juliet *wouldn't notice* her . (Luisa, file 47)

[147] But if they do bad things to other migrants no one *would forget* it. (Luisa, file 81).

In 83% of the Will-Future contexts she creates , Luisa uses the form for prediction-based reference [148-9] . Until Period 4, at the end of the first year of the study, this is the only use.

[148] I'm sure that Dazza *will pick up* the pingeon [: pigeon]. (Luisa, file 7)

[149] You*'ll never forget* that expirence [: experience]! (Luisa, file 21)

From Period 4 on, the *will* form is also used for intentions and plans. Initially this is because *be going to* has not emerged in her interlanguage, but even after the emergence of the Go-Future *will* is still the most frequent form used by Luisa for intention-based reference. However in the contexts she creates it would probably be the form preferred by L1 speakers. From Period 9 on there are 15 'intention' contexts. 9 of them are introduced by verbs like *decide* and *promise* and she uses what Leech describes as a 'performative' *will* rather than a 'premeditative' *be going to* to reinforce the feeling that in the act of speaking, the decision is made [150-2]. Leech says that this use of *will* could be considered a 'volitionally coloured future '(2004, ,86).

[150] She said *she'll be* back tomorrow. (Luisa, file 68)

[151] Hamlet decides that he *will act* like a mad man from now on. (Luisa, file 86)

[152] *I'll come back* as quickly as I can Ann, I promise. (Luisa, file 68)

Luisa only overgeneralizes the *will* form 4 times, in each case in a Subordinate Future context – a dependent clause where the Simple Present form would be preferred by L1 speakers:

[153] In the following day if the bird *won't feel* better he'll take it to the a vet. ... At last when Cherokee *will be* health and strong again it will be relly [: really] hard for them to split up. (Luisa, file 7)

will +progressive infinitive

Luisa also uses *will* with the progressive infinitive 5 times. Four instances are typical of the form's use, referring to temporary or ongoing situations in the future:

[154] You'll have a good fun and meet some new people . *You'll soon be going out* with them and won't stay at home .

(Luisa, file 41, advising a girl to join some clubs so that she can make friends)

However one use, in Period 16 towards the end of the study, is an instance of what Leech (2004) calls the 'future as a matter of course' , a use which he says appears to combine the 'prediction' meaning of *will* with the 'arrangement' meaning of the Futurate Present Progressive, and which may have developed as a means of avoiding the overtones of volition which *will* sometimes combines with its prediction meaning (see 3.6.2). The sense here is that since the famous footballer will be coming to the area anyway, he will meet a local reporter to do an interview:

[155] Francis O'Connell *will be coming* for the finals of Junior Championship ..... and it seems like he is interested in the interview. (Luisa, file 97)

Leech notes that this form is generally used to refer to the near, but not too immediate future, and this is also the case here.

Sara

Like Luisa, Sara appears to have acquired the Will-Future before the start of the study. She creates 89 contexts for it in her data, using it accurately 25 times until Period 4, after which she once fails to supply a *will* form [156] and on 4 occasions uses an inaccurate form, twice supplying a *will*+Simple Present instead of *will*+infinitive in contexts where the Simple Present would also have been appropriate, suggesting a momentary confusion between the two options [157-8].

[156] The house will be beside a lake and it will locate on its own. I *let* my best friend to live with me. My family and her family, we will work together. (Sara, file 14)

[157] It is the one of the best poem I've ever read . It has certain meanings that *will makes* us understand the world around us . (Sara, file 45)

[158] Then he gets more provoked and he *will tries* to take revenge. (Sara, file 65)

Sara uses the Will-Future for prediction-based reference 79% of the time [159] From Period 2 on she uses the form for plans and intentions as well [160].

[159] It will interfere with the development of responsible behaviours. (Sara, file 67)

[160] The driver John said that he will follow us when the bus is repaired. (Sara, file 34)

Sara mainly uses *will* with 3rd person subjects – 76% in the case of prediction-based uses, and 68% for intention-based ones. And in her case, prediction-based uses are as likely to refer to an imminent future (today, tomorrow) as intention- based ones are.

Sara does not overgeneralize the *will* form.

Ana

Just before the school holidays preceding this study, Ana wrote about her summer plans, and as this excerpt from the text shows, at that stage she was using base verb forms and *to be* in future contexts, and overusing *will* with modals :

[161] When finish this first year I *will must come back* to Lithuania for two week [...] but I think my holiday quicly *finish* and I *be* a happy. I watched news and said in Ireland summer *to be* very warm. (Ana file 3)

In the early months of the study Ana creates only two contexts for grammaticalized future forms in her written data and in both cases fails to use a future form :

[162] Then I said for him that I can to ride the bicycle and *don't fdt* again (Ana, file 12)

[163] But godmother tells that you [/] this all goods on midnight *it's been* the same what's at starts.
(Ana file 13, recounting the Fairy Godmother's warning to Cinderella).

In the initial stages Ana sometimes relies on lexically future verbs such as *want, have to* and *like,* or produces present tense narratives and responses to prompts which seemed to call for the future. But from Period 4 on she creates 80 potential contexts for the Will-Future and only once fails to supply an accurate form. Initially she employs the *will* form to express both prediction-based [164] and intention-based meanings [165] . Prediction accounts for 82% of all her Will-Future contexts and from Period 8, half way through the study, it is, with one exception, the only meaning associated with the will form.

[164] Of course you will find work but salaries will be different and smaller than others who finished university. (Ana, file 41)

[165] And she told me that we [//] on Monday we would go with [//] sailing with her. (Ana, file 34)

Ana overgeneralizes the *will* form 8 times in Subordinate Future contexts, in fact it is the form most frequently used by her in these dependent clauses (8/14) where an L1 speaker would prefer a base/Simple Present form:

[166] You have to have a good family and good condition for yourself and them. So if *they'll be* happy, you'll be auto [/] happy too. (Ana, file 41).

8.3.3 Summary

Research suggests that *will* +infinitive is the first grammaticalized future form to emerge (3.6.2) and this is the case for the three learners in this study. The form is used with a very high degree of accuracy by Luisa and Sara right from the beginning and it is probable that they had acquired it before the start of the study. However the form does not appear in Ana's written data until Period 4. Once it does though, it is always supplied where a potential context has been created, suggesting that, in Ana's case at least, the form is quickly and firmly entrenched once it emerges.

Prediction and Intention

The Will-Future is most typically associated with prediction (3.6.2) and this meaning accounts for around 80% of the learners' overall use of the form. In fact *will* is more or less the learners' sole means of expressing prediction until around half way through the study, when other emerging forms become available (Period 8 for Sara, Period 9 for Luisa and Ana). Up to this point Luisa uses *will* in 61 out of 63 prediction contexts, Sara in 28 out of 29, and Ana in 28 out of 30.

In the absence of the forms more typically associated with intention-based reference, the Will-Future is also the main means (and in Ana's case, the sole means) of talking about plans, intentions and resolutions. Once the Futurate Present Progressive and the Go Future emerge, *will* can move into its more prototypical role in the learners' interlanguage, and this will be discussed later in Chapter 9

Imminence

The Will-Future can be used to express both a near and a distant future (3.6.2) and in this data it is used to express imminence [167-9] as well as remoteness, especially during the time that it is the main form used for intentions as well as predictions. Once the Go-Future emerges there is a noticeable change in this aspect of the will form's use, and this will be discussed later in Chapter 9

[167] But we won't be able to get there by car.
(Luisa, file 38, referring to that evening)

[168] I was waiting Miss Hannon that she will correct my test.
(Ana, file 18)

[169] I am sure we will beat them tomorrow. (Sara, file 50)

Grammatical Person

Dahl (2000a) reports that in French and Spanish the inflectional future tense whose typical meaning is prediction is more likely to be found with 3rd person subjects, whereas the de-andative (*go*) form with its typical meaning of intention is more likely to be found with 1st person subjects.

Bardovi-Harlig (2004b) notes that her learners used a higher proportion of third person subjects with the Will-Future than the Go-Future. In this data around two thirds of *will* forms are found with 3rd person subjects. As might be expected, a higher percentage is found where *will* is used for prediction-based meanings, than when it is used for intention-based meanings where the learners are more likely to be discussing personal plans and intentions.

Overgeneralization of the Will-Future

While there may be a loss of precision, of subtlety of meaning, when it is employed instead of another form, *will* + infinitive is acceptable in many, if not most, future contexts (see 8.7 for a discussion). However dependent clauses introduced by conditional, temporal and manner conjunctions (*if, unless, when* etc) do not provide a 'friendly' context for *will,* as here L1 speakers prefer the Simple Present (3.6.2) Both Luisa and Ana overgeneralise *will* to these dependent clauses – Luisa does so initially and Ana continues to do so until late in the study. This overuse could be taken as an example of *will* functioning in the learners' interlanguage as a general marker of the future, as a default form supplied in clauses identified as having a future meaning. Luisa and Ana's developing use of the Simple Present in these dependent clauses - and Sara's very accurate use – is discussed in the next section.

8.4 Simple Present forms in future contexts.

8.4.1 Introduction

According to Leech (2004) the Present Tense in a broad sense encompasses both present and future domains of time and this helps to explain why English, which does not have a future tense as such, uses present tense or present tense auxiliaries (*will, shall, is going to, is about to*) to express future time. This section looks at the learners' developing use of the Simple Present in future contexts, starting with the more frequently occurring Subordinate Future.

8.4.2 The Subordinate Future

After *will* + infinitive the second most frequent future construction in English is Simple Present as Subordinate Future (3.6.2). It is also the second most frequent in this data, occurring most often – about half the time - in dependent clauses introduced by *if*, but also quite often in clauses following temporal conjunctions.

Luisa creates 24 Subordinate Future contexts. There are three in Period 1, with two overuses of *will* and one base form. She does not create another context until Period 6, but from here on she uses the Simple Present form appropriately most of the time [170-1] – there are just 2 further instances of the overgeneralisation of *will.*

[170] If you trust your chi [/] son or daughter, he or she will give you the same in return. (Luisa, file 82)

[171] They are going to make their own decisions weather [: whether] you let them or not. (Luisa, file 82)

Sara appears to have acquired this construction before the start of the study. She uses the Subordinate Future 32 times from Period 1 on, always accurately and appropriately :

[172] If they own up the crime they'll be exiled from the city. (Sara, file 55)

[173] It will a [/] make you a huge difference when you have freedom. (Sara, file 48)

Ana struggles with this construction. She creates 14 contexts, the first in Period 5, and overuses *will* 8 times. It is only towards the end of the study, from Period 15, that her use is generally accurate:

[174] If you get education you will get better job. (Ana, file 79)

8.4.3 The Futurate Simple Present

The Simple Present with future reference is only infrequently used in independent clauses in English (3.6.2) and is normally accompanied by an adverbial establishing the future meaning. According to Leech (2004) it

can represent either a 'future assumed to be fact' or ' a plan or arrangement regarded as unalterable'. Luisa creates only 4 contexts in the data, the first in Period 7, and Sara only one, in Period 12. Both use it appropriately :

[175] The bus *is* at five past seven. (Luisa, file 38)

[176] Next Monday *is* the operation of Anna's mum. (Sara, file 56)

Ana creates one context - in Period 4 - but does not use the Simple Present; instead *will* functions as a marker of the future. Ana is on the athletics team and her coach has told her that there is a competition at Santry Stadium the next day:

[177] Then I met Miss who told me that tommorow [: tomorrow] *will* Santry. (Ana, file 18)

8.5 The Futurate Present Progressive

8.5.1 Introduction

The Present Progressive form can be used in English to refer to 'a future event anticipated by virtue of a present plan, programme or arrangement' as distinct from a future intention (Leech 2004, 61). Mainly confined to contexts for this specialized meaning, the form is much less frequently used than either the Will-Future or the Go-Future (3.6.2). In this data the Futurate Present Progressive is employed in just under 7% of the future contexts created by the learners (which is higher than the 2% recorded in Bardovi-Harlig's 2004 study).

8.5.2 Emergence and developing use

All three learners use the form once early on, but then it does not appear again until about half way through the study - Period 7 in the case of Luisa and Ana, and Period 8 in the case of Sara. For all three it is a late emerging form (though emerging before the Go-Future for Luisa and Ana, and at around the same time for Sara).

Luisa uses the Futurate Present Progressive 12 times, Sara uses it 7 times and Ana 11. As is the case in L1 patterns of use (Section 3.6.2) the form mainly occurs with ' doing' verbs involving conscious human agency, and in the data it occurs 57% of the time with *go,* and 20% with *come.*

Leech (2004) notes that the difference between 'arrangement' and 'intention' is a slight one and yet in this respect the Futurate Present Progressive is almost always used very precisely by the learners , referring to a situation which has been planned, and often to a social arrangement already made:

[178] *She's going to* theatre with her friends. (Luisa, file 38)

[179] There are only few weeks before your Leaving Cert exams *are starting*. (Luisa, file 96)

[180] This will be my last holidays with them because I *am going* to Canada. (Sara, file 50)

[181] As I told you yesterday *we're going* for hiking and *I'm taking* you with me. (Sara, file 54)

182] A famous writer Emily Dickinson *is coming* to Dublin to publish her new poems. (Ana, file 87)

[183] On Monday morning she rang me and told me that we would meet near shop and that her boyfriend *is coming*. (Ana, file 34)

There are only two occasions where the Futurate Present Progressive is used in a context where it is referring to an intention rather than an arrangement:

[184] And when they told me they go to South America, I said *'I'm not going* with you this time'... what surprised me was that they agreed . (Luisa, file 65)

[185] 'That's it!' I thought. *'I'm going* home because I will be in trouble.' (Ana, file 64)

The Go-Future is the form most typically associated with intention, but this use of the progressive, expressing an intention to *go* somewhere, occurs frequently enough to be noted in pedagogical grammars as an alternative use (Parrott 2010, 202). According to Dahl (2000a) there

may be a tendency in English to avoid *be going to* with movement verbs, favouring the use of the progressive instead, and that may be what is happening here, with the learners avoiding a *'going to go'* construction or conflating the second *go* form with the first. In any case there is no instance in the data of the Go-Future being used with the verb *go (*or with *come).*

In summary, the Futurate Present Progressive emerges relatively late in the learners' interlanguage but is used appropriately from its first appearance and almost always with its specific 'arrangement' meaning. The effect of its emergence on use of the Will-Future, and its relationship with the emerging Go-Future will be discussed in Chapter 9.

8.6 The Go-Future

8.6.1 Introduction

One of the two major grammaticalized future forms, the Go-Future is a relatively recent arrival in the English language (the spread of the form only dating from the 17thcentury) and while it is becoming increasingly common in informal English, it is still only used about half as often as the Will-Future is (3.6.2). According to Leech (2004) the form is most typically used to express a future linked to the present – either the future outcome of present intention or the future outcome of present cause. However he notes that in informal English it is also increasingly used in a more general way as a more neutral future auxiliary and all three learners in this study do use *be going to* in this way, though in the case of Luisa and Sara the form is mainly used to express the more typical 'present intention' and 'present cause' meanings.

8.6.2 Emergence and developing use

L2 Acquisition research has found that *be going to* emerges after *will* in the interlanguage of learners, emergence being defined as the appropriate use of the form with three different verbs (Bardovi-Harlig, 2004b). This is the case here – in fact for Luisa, Sara and Ana *be going to* emerges quite late, with *will* dominating their future expression for most of the first two years of the study.

Luisa

Luisa only creates 11 contexts for the Go-Future, fewer than Sara or Ana. She first uses the form at the start of the study, in Period 1:

[186] But I don't think I'm going to have childrens [/] children in the future. (Luisa, file 4b)

But then she does not use it again (apart from one prompted use) until Period 9, in the meantime rarely creating a context for intention-based future reference, and when she does, using the Will-Future to express it. *be going to* is always used accurately, occurs with 7 different verbs and is employed to express general predictions [187] and plans [188] as well as the more typical Go-Future meanings of intention [189] and contingent future (future of present cause) [190]. She also uses the form with reference to a remote future [191] although most of her use refers to an imminent situation or event.

[187] things *are going to be* better. (Luisa, file 81)

[188] This holiday has only started and *I'm going to spend* it with my mum. (Luisa, file 68, speaking about an existing arrangement)

[189] *I'm going to make* them sorry, no mather [: matter] what! (Luisa, file 85)

[190] I knew that *I'm going to be* the oldest in the class because I had to repeat one year because of my English. (Luisa, file 76, talking about how she felt arriving at St John's on her first day)

[191] *I'm* gonna [/] *going to design* and build bikes some day. (Luisa, file 77).

So although Luisa only employs the Go-Future 11 times, she uses it with the full range of appropriate meanings.

Sara

Sara creates 21 contexts for the Go-Future. In Period I she uses a hybrid form when responding to the question *What is Dazza going to do next?:*

[192] I think he *will going* to take the bird home. (Sara, file 5)

In Period 4 she uses *be going to* for the first time, but in a context where it could also have the '*go* +infinitive of purpose' meaning from which the grammaticalized future form originally developed. She is describing how, as a small child in India, she saw a snake in the garden and thought it was harmless. Her father came looking for her:

[193] He saw that I am going to [//] *was going to touch* it. He killed the snake. (Sara, file 16)

Her next use is months later, in Period 7. Then in Period 8, she creates 18 future contexts and uses *be going to* in 9. Apart from one general prediction her use is equally divided between the typical Go-Future meanings of present intention [194-5] and the future of present cause [196].

[194] I *am going to* chose [/] *choose* Rita because she seems to be mature (Sara, file 37)

[195] In the extract he said that he *is going to try* acting. (Sara, file 43)

[196] I was a hundred percent sure that we *are not going to survive* this trip.... The thought was torturing me, me and my friends *were going to die.*
(Sara, file 32, referring to the dangerous situation they were in)

From Period 9 on, the form is mostly used to express intentions, but also plans:

[197] We *are going to split* in teams. (Sara, file 54)

Sara mostly uses the form to refer to the near future, but there are several examples of a remoter (but not very remote) time reference:

[198] Next month we are going to college so everyone [/] we *are* all *going to split.* (Sara, file 34)

Ana

Ana creates 18 contexts for the Go-Future. Like Sara's, her first use of *be going to*, in Period 5, is slightly ambiguous, as it could be a description of motion, using *going* in its literal sense; or it could be a prediction based on what is happening at the time. Looking at a picture of a rocket leaving a launch pad, she says:

[199] I think this machine is going to fly to a space. (Ana, file 24).

Ana does not use the form again, unprompted, until Period 9, and again the two uses here involve motion towards a goal, and out of context could be considered examples of the '*go*+infinitive of purpose' construction out of which the grammaticalized future form developed :

[200] She said me 'Hang on, I'm going to get the ball'. (Ana, file 56)

[201] So I decided that I'm going to visit my granny. (Ana, file 61)

After this, the Go-Future does not appear again in Ana's data with a verb of motion - *be going to* is unambiguously associated with future outcomes.

Although her earliest uses of the form refer to plans or intentions, Ana differs from the other two learners in that she mainly uses *be going to* for predictions (12/18) and these are mostly the 'neutral' future of prediction (rather than the future of present cause). As already mentioned above (8.6.1), in informal L1 English the Go-Future is increasingly employed as a fairly neutral future auxiliary. However in Ana's case her use may be due to her strong tendency to overgeneralise or overextend an emerging form for a short period of time; this tendency has already been noted in Chapter 7 in relation to the Present Perfect, where she used the form 11 times in Period 8, but only 3 of those uses were in a Present Perfect context (7.4.4)

Something similar may be happening with her use of the Go Future in Period 13. Before, and after this period, *will* dominates her future talk but in Period 13 she uses *be going to* 9 times, and the *will* form only 7 times even though all the contexts are for 'neutral' prediction, a natural

environment for the Will-Future. However there is some evidence of the two forms being used contrastively. The Go-Future is typically associated with imminence and there are 6 near future (now, today) contexts in Period 13 and *be going to* is used in all 6. For example:

[202] So here I am, standing and waiting [/] wondering what is going to happen today. (Ana, file 66)

[203] Am I going to find any friends today? (Ana, file 66)

[204] Maybe it's going to be the best idea to write about how beautiful she was. (Ana, file 68)

And there is one instance where she originally used the Go-Future to talk about a remoter future [205] , but then changed her mind and rephrased, using the Will-Future [206] .

[205] He told me that *I'm going to be* a great writer in a few years later from now.

[206] He told me that ' *you will be* a great writer in a few years later from now'. (Ana, file 67)

Ana is the only one of the learners to use *be going to* in atypical contexts; there are two instances, one in Period 12 [207] and one in Period 15 [208] which might suggest that even towards the end of the study she is still in the process of sorting out the Go-Future's form-meaning associations.

[207] We promised ourselves that we are going to [//] were going to live in this house when we *were going to turn* eighteen . (Ana, file 65)

[208] I started to study for my English test which is [/] *was going to be* the next day. (Ana, file 79)

The 'be about to' construction

be about to is a minor future form, occurring infrequently in L1 English, however it appears in the data for all three learners soon after the emergence of the Go-Future. According to Leech (2004) *be about to* is close to the meaning of *be going to*, except that it suggests greater immediacy. Sara uses the form 4 times, Luisa 3 times and Ana once, and in all cases it is used with reference to an immediate future:

[209] And he thought that his mother is in the stable at Bethlehem where baby Jesus *is about to be born* . (Ana, file 62, discussing a poem about Christmas morning)

[210] I was too busy chatting with my friends in Yahoo. I signo [/] *was about to signout* when I saw an unopened email from two weeks ago . (Sara, file 50)

[211] Even when his shoe *was* breaking [/] *about to fail* him he had a plan... even if it ment [: meant] him crawling his way forwards on the ice-cold snow .
(Luisa, file 91, talking about Nando Parrado's ordeal in the Andes)

8.6.3 Summary

The Go-Future emerges late in the learners' written samples, but once it emerges it is used accurately – the correct form is always supplied – and in the main it is used to express meanings typically associated with the form.

Imminence

The 'future of present intention' is a very typical use of the Go Future in L1 English (3.6.2). It accounts for half the *be going to* constructions in the learners' data, and the intention almost always refers to a near future (today, tomorrow). In the case of Luisa and Sara, when the form is used for predictions they are mainly predictions based on present circumstances, in other words the 'future of present cause' (or Contingent Future) which has also been identified as a typical use of the construction (Leech, 2004; Radden and Dirven, 2007). While two thirds of Ana's predictions are of the 'neutral' type, most refer to an immediate future (today, tomorrow) in keeping with the form's typical association with imminence. In fact the feature which characterizes all three learners' use of the form is that it is mainly used in reference to the near future (and this will be discussed further in Chapter 9).

Formulaic Use

Five of Bardovi-Harlig's 16 learners exhibited some formulaic use of *be going to* (for discussion, see 3.6.3) . While formulas may have played a role in Luisa, Sara and Ana's early conversational use of the Go-Future, there is no evidence in the written data of any formulaic use of the form,

which is found with a wide variety of verbs . In fact 32 different verbs are used in the 50 contexts created by the learners (Figure 8.3)

ask be brush choose design die do enjoy find fly forget get graduate happen have learn like live make move pay say sleep spend split surprise survive take touch try turn visit

Figure 8.3 Verbs used with the Go-Future

Grammatical Person

The Go-Future is found with 1st, 2nd and 3rd person subjects. And there is a marked difference here, between intention-based and prediction-based reference. When the learners use the form for intentions and plans, 1st person subjects predominate (20/24) but when they use it for predictions, the form is most frequently found with 3rd person subjects (16/24). Bardovi-Harlig (2004b) says that the study of grammatical person was 'somewhat thwarted' in her corpus as the learners mainly wrote about themselves , though she observed that scheduling and prediction, when they appear, allow an opening for the 3rd person.

Luisa, Sara and Ana's use is in line with the association of 1st person with intentions, and 3rd person with predictions found for the Western European languages reported on by Dahl (2000a) and with Romer's (2005) observation that in the BNC and Bank of English corpora a large number of 1st person subjects occur with *be going to* when it is used to express intentions .

8.7 The dominance of the Will-Future & the late emergence of the Go-Future.

Of the two main grammaticalized future forms in English, *will*+infinitive is the most frequently occurring. In spoken British English it outnumbers *be going* to by a ratio of 2.5:1 and in American English by a ratio of 1.6:1 (Biber et al., 1999). In Bardovi-Harlig's written sample the Will-Future outnumbers the Go-Future by a ratio of 5.8:1 (1407:242) suggesting, she says ' the degree to which *going to* is underrepresented in the interlanguage of these learners'(2004b, 182). In Luisa, Sara and Ana's data *will* outnumbers *be going to* by a ratio of 6:1 (302:50) which is very similar to that found in Bardovi-Harlig's study.

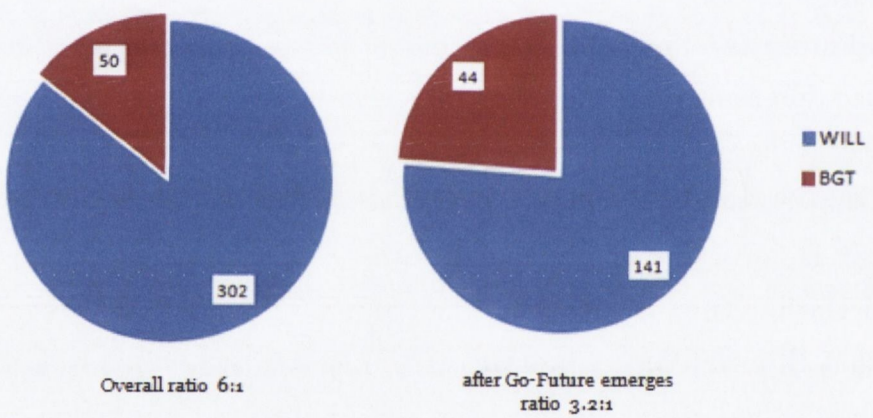


Figure 8.4 Will-Future use v Go-Future use

From the period in which the Go-Future emerges (Period 8 in Sara's case, and Period 9 in Luisa's and Ana's) the ratio is much closer to that found in L1 English – it is 3.2:1 in favour of *will* (Figure 8.4) This ratio disguises the individual differences however, with Sara and Ana, at 2.5:1, and 2.7:1 respectively and Luisa at 5.5:1. As already noted (8.3.2) most of the intention-based contexts which Luisa creates from Period 9 on involve decisions or promises and , as an L1 speaker might, she uses the 'performative' *will* rather than the 'premeditative' *be going to* in all of

these contexts; this accounts, in part, for her higher ratio of Will-Future to Go-Future.

Why does *be going to* emerge so late in the data, relative to *will*, when it is quite frequent in spoken English and as a result probably quite frequent in the input the learners are exposed to? Research has shown that learners use lexical means before morphological means to express temporal relations (for a discussion see 3.4) , and Bardovi Harlig (2004a) considers the possibility that her learners may perceive *will* as a lexical expression, so giving it an advantage over *be going to*. Among the facts which she cites in support of this possibility are that *will* is a free morpheme and that it virtually lacks variants in her learner corpus (where there are very few tokens of the phonologically reduced form, used by only two learners). However Luisa, Sara and Ana all regularly use variants of *will* from early on in the study - *will, 'll, won't* and, in backshift contexts, *would, 'd and wouldn't* . *Will* emerges in Period 4 for Ana and by Period 5 she is using the contracted form with pronouns. In fact the majority of Luisa and Ana's use with pronouns involves the contracted form, and 20% of Sara's does. Ana's data suggests that even if *will* is initially perceived as a lexical item (she does use it twice, early on, without the infinitive) this perception does not last long and cannot explain *will*'s long period of dominance.

Compared to *will* +infinitive, *be going to* +infinitive is a complex form inflected for person, number and tense on the auxiliary *be*. Bardovi Harlig briefly considers formal complexity as a factor but concludes that it is unlikely to explain *will*'s dominance. The data in this study supports that view, as the learners early on demonstrate an ability to handle compositional tense-aspect forms. The progressive form is used appropriately and accurately by Luisa and Sara right from the beginning of the study, and by Ana from Period 3. And right from the beginning, all 3 learners are able to produce constructions like the infinitive of purpose. In any case there is no real evidence in this written data of the learners

struggling with the Go-Future form. They use it accurately right from first use (or in Sara's case, second use). Luisa actually first uses *be going to* early on in Period 1 but then does not employ it again, unprompted, until Period 9, in the mean time having used *will* 77 times in future contexts.

One possible reason for the delay in the emergence of *be going to* may be that *will* not only emerges first as a prototypical future, but is a particularly useful prototype for learners because it is so often an acceptable alternative to other future forms. There may be a loss of precision, of subtlety of meaning, when *will* is used rather than another form, but still *will* does the job, it conveys the basic meaning (in a way that the Simple Past, for example, cannot routinely substitute for the Past Progressive or Present Perfect without a change in the basic meaning). As the following examples illustrate, *will* functions adequately in a wide variety of future contexts:

They arrive on Wednesday.
They will arrive on Wednesday.

He is spending Christmas with us.
He will spend Christmas with us.

She will be expecting to find us here waiting when she arrives.
She will expect to find us here waiting when she arrives.

They are going to be late for the meeting.
They will be late for the meeting

So given how useful *will* is as a general-purpose future, how and why does *be going to* break through? It is possible that the emergence of the Go-Future is driven by the learners' need or desire to use grammatical, rather than lexical or contextual means, to 'fine tune' their future expression and is facilitated by their having reached a stage of language development where this is achievable – where in fact they have already begun 'fine-tuning' other areas of their temporal expression. This will be discussed further in Chapter 9.

8.8 Conclusion

An analysis of the learners' written data reveals that the development of future time expression is a gradual process which continues over the four years of the study, with the emergence of new constructions affecting existing patterns of use, as new form-meaning associations are forged.

will +infinitive is the most frequently occurring future form in English and is the first to emerge in the learners' interlanguage. It is used from the start by Luisa and Sara but does not appear in Ana's data until Period 4, near the end of the first year of the study. Until Period 8 *will* is the only form used to express predictions and by far the most frequent form used for intentions; it is also employed by Luisa and Ana as a subordinate future in dependent clauses.

The Simple Present is used appropriately in Subordinate Future contexts from the beginning by Sara, and by Luisa from Period 6; however its use does not become established in Ana's temporal system until very late in the study at Period 15. For Luisa and Ana the Futurate Present Progressive becomes the main means of expressing plans and arrangements from Period 7 on, and for Sara from Period 8.

will loses its monopoly in prediction contexts when the Go-Future emerges (just after the Progressive in the case of Luisa and Ana, and in Sara's case around the same time as the Progressive). *be going to* is used for both prediction and intention-based reference from the start, and in both contexts equips the learners with a grammatical means of explicitly linking the future to the present. The *'be about to'* construction appears after the Go-Future, adding to all three learners' repertoires the means of expressing the immediacy of a future event.

For the first half of the study the learners' future time talk is completely dominated by *will*. Gradually other forms appear, or an existing form like the progressive moves into a new environment, adding future reference to

its other more typical roles. By the end of the study, while *will* +infinitive is still the most frequently used, the learners, between them, have employed six different forms and used them to express a range of meanings from the general 'neutral future of prediction' to the very specific 'matter of course future'.

9 Summary and Discussion of Main Findings

9.1 Introduction

The main focus of this study is an analysis of the L2 acquisition of past and future temporal expression. Chapter 7 presented a detailed picture of the progress made by the learners in acquiring past-related tense-aspect forms, and Chapter 8 detailed the progress they made in relation to the acquisition of future temporality. Some background information on vocabulary development, lexical diversity and overall accuracy of language use was presented in Chapter 6.

In this chapter the principal findings of the study are summarized and then discussed in relation to the research questions set out in 4.3 and the literature outlined in chapters 2 and 3. Section 9.2 looks at aspects of the learners' overall language use and considers what evidence there is of continuing development over the four years of the study, with particular reference to lexical diversity and accuracy of language use. Then past-related (9.3) and future time expression (9.4) are discussed, both in relation to the orders of emergence established in L2 acquisition research, and in relation to the learners' developing use of temporal expression – the spread of existing forms from prototypical to less typical uses for example , or the integration of new forms into the system. Section 9.5 provides an overview of the order of emergence of form-meaning associations in the temporal system as a whole. In usage-based models of L2 acquisition, input and interaction are regarded as central to language learning, with this learning affected by a range of frequency-related factors. Section 9.6 discusses what evidence there is in the learners' data that their acquisition of tense-aspect forms and other constructions is influenced by such factors as frequency, frequency distribution, salience and prototypicality.

9.2 Lexical development and accuracy of language use

There is clear evidence in the data that Luisa, Sara and Ana make significant progress in acquiring English and that they continue to make progress throughout the four years of the study. While individual differences are evident in the rates of progress made in different areas of the system, for all three learners the acquisition of English is a gradual process, as research suggests it would be (Klein et al., 1995; Ellis and Larsen Freeman, 2006).

When they enrolled in First Year in St John's , Sara had already been in Ireland for one year, but Luisa and Ana had just arrived in the country and initial assessements showed that they were at an earlier stage of English language development than Sara (Section 4.6.2). Luisa appeared to progress particularly quickly through the early stages of acquisition however, and by the beginning of Second Year (and the start of this study) the gap between her and Sara had narrowed.

9.2.1 Lexical development

The range and variety, or diversity, of vocabulary deployed by a speaker or writer is considered a useful predictor of overall language proficiency (Jarvis, 2013). To obtain a general picture of the development of proficiency an analysis was conducted using the *vocd* program (Malvern and Richards, 2002). Research suggests that lexical diversity (*D)* as measured by this program, is the best predictor of lexical proficiency currently available (Crossley et al., 2011).

The analysis shows that Sara and Ana begin the study with mean D scores lower than those found for the most proficient 7 year old L1 writers in Malvern et al (2004), and end the study with mean D scores comparable to those Duran et al. (2004) reported for adult L1 writers (Table 6.3, Figure 6.1). Sara's lexical diversity increases significantly over the first

two years of the study and she makes steady – always upwards - progress after that, obtaining her highest mean D score at the end. Ana's lexical diversity increases slowly at first and then there is a significant rise in the first half of the second year and again in the fourth year, with a period of stagnation in between (her minimum D values continue to rise steadily in each period however). Ana also achieves her highest mean D score at the end.

From the start of the study Luisa's writing displays a high degree of lexical diversity. Her D score in Period 1 is 81.43, compared to 51.73 for Sara and 48.76 for Ana. By the middle of the second year Luisa has achieved a mean D score comparable to an adult L1 writer (90.82), and while there are small fluctuations in the D after this, it never drops below adult L1 level.

So for Sara and Ana the development of lexical diversity continues for the whole four years of the study, with them ultimately achieving a nativelike standard, while Luisa achieves this standard in the second year of the study.

While Sara and Ana did not continue to study their L1 after arriving in Ireland, Luisa attended regular Polish language and literature classes and was considered a proficient writer by her Polish teacher (as discussed in Section 4.6.2). It is possible that her strong – and still developing- literacy skills in her L1 gave her an advantage over the other two learners when it came to the development of lexical proficiency. The Linguistic Interdependence Hypothesis proposed by Cummins (2000) suggests that literacy skills developed in the L1 will transfer to the L2. Luisa may, for example, have been able to transfer to English the writing skills she had developed in Polish, so deploying her English vocabulary more effectively. It is also possible that her continuing L1 lexical development supported her L2 lexical development. For example, quite a few of the less common words which appear in her written data are very similar to their Polish

equivalents and may have been acquired first in Polish, facilitating their acquisition in English (examples are given in 6.2.3).

For all three learners the data provides evidence of the continuing expansion of vocabulary, with the words appearing for the first time in the final year of the study tending to be formal or technical words (see 6.2.2. for examples) which are relatively infrequent in the BNC corpus of L1 English (this will be discussed further in Section 9.6). There is also evidence of the learners' vocabularies continuing to grow in depth as well as breadth, with polysemic words occurring with new meaning associations right up to the end of the study, with verbs such as *take put, make, know* and *get,* for example, used in new senses in the last month of the final year.

An analysis of the learners' use of vocabulary reveals patterns very similar to those found for L1 speakers of English. For example, the 20 most frequent verbs in the BNC corpus (Leech et al., 2001) account for around 70% of all verb use in the data. These verbs are likely to be very frequent in the input which the learners received and usage-based theories of language learning suggest that this would have facilitated early acquisition (Beckner et al., 2009). Their continuing frequency of occurrence into the later stages of language development however is perhaps more likely to be due to the fact that these verbs are generally very useful in communication, many being polysemous as outlined in Crossley et al. (2010). Therefore Luisa, Sara and Ana probably employ them as often as they do for the same reasons L1 speakers do.

9.2.2 Accuracy of language use

While an accuracy analysis, on its own, could be considered a very negative way of looking at a learner's interlanguage (Devitt, 1992), such an analysis can contribute something to the overall picture of language development (Ellis and Barkhuizen, 2005) – especially if it is based on relatively spontaneously produced data, where the learners are more likely

to be employing their implicit, rather than their explicit linguistic knowledge (6.3.2). *errors per 100 words* is a general measure of accuracy which has been widely used by researchers (5.4.2) and it is the measure employed here (while spelling, punctuation and other orthographic errors are often included in error analyses, they are not included in this analysis).

All three learners achieve improvements in their general accuracy of language use over the course of the four years (Table 6.6, Figure 6.2). Progress is sometimes slow however – due in part to the emergence of new forms and constructions which have to be integrated into the learners' existing language system (Dietrich et al., 1995) and there are periods where there is no reduction, or even a slight rise, in the *errors per 100 words* rate .

In Period 1 Ana's rate, at 16.7, is more than double that of Luisa and Sara (7.5 and 7.3 respectively) – a reflection perhaps of the fact that she was at an earlier stage of language development than the other two. Her overall accuracy improves markedly during the next year and a half, with the error rate dropping to 7.2, but then there is no further improvement until the fourth year when the error rate drops to 6 in Period 7, before rising slightly to 6.4 in Period 8. Luisa's error rate halves over the first two years, after which gains in accuracy come slowly, but by the end of the study she has achieved a very low error rate of 2.4 *errors per 100 words*. Sara makes the least progress of the three, with some areas – such as verbal agreement of person and number - remaining problematic for her. There is very little change over the four years, though she achieves her lowest error rates in Periods 7 and 8 at the end of the study.

Some developmental sequences are traversed by all three learners, and accurate use achieved in these areas, before the final year of the study. For example, Luisa is using the English negation system accurately over 90% of the time right from the start, Sara does so from Period 4, at the

end of the second year, and Ana from Period 6. However while Ana achieves accurate use of do-questions in Period 3 and Luisa in Period 4, it is only in the fourth year of the study that Sara's use is mostly accurate. Although she enters the final stages of interrogative development in Period 2 (producing questions with the auxiliary in second position after a wh-word) Sara routinely marks both the auxiliary and the main verb for tense until Period 7.

It seems to require a considerable length of time for some aspects of the English lexical system to be acquired. At the end of the study, for example, the three learners are still having difficulty sorting out appropriate use of frequently occurring prepositions. The article system can also be problematic for those whose L1 does not have articles (Luk and Shirai, 2009) and it accounts for between a fifth and a quarter of all the errors made by the learners. Once an accuracy rate of around 80% has been achieved, further progress seems to be slow, possibly because it is the more complex uses which remain to be acquired. Sara's accuracy rate for article use is 79% in Period 1, and Luisa's rises from 55% in Period 1 to 81% in Period 2. Both achieve their highest accuracy rate – 92% - in Period 7 and at this stage very few errors relate to the more typical contexts for article use. In Period 1, after a year in an English language environment, Ana's article use is only accurate 36% of the time. By Period 3, a year later, the accuracy rate has risen to 60% and by Period 7, a further two years later, to 79%. At this stage she is still making some relatively simple errors. For all three learners article use is most accurate in the final year, suggesting that, although progress may be relatively slow, it is still continuing to be made.

Overall, the three learners achieve their highest level of accurate use of the language in the final year of the study (Figure 6.2). In Sara and Ana's case, a marked drop in error rates in Period 7, is followed by a rise in Period 8. However Sara's error rate did fluctuate twice before this, and Ana's did so once before, and the final *error per 100 words* rate in the

fourth year, that for Period 8, is lower than those achieved in previous years.

9.2.3 Conclusion

In summary, the evidence relating to the development of lexical proficiency, and the achievement of overall accuracy of language use, suggests that while individual rates of progress may differ, English language development does continue for all three learners for the four years covered by the study. All three learners eventually achieve a level of lexical diversity comparable to an adult L1 writer; Luisa does so in the second year and Sara and Ana at the end of the study. Overall accuracy of use continues to increase, with Luisa in particular ultimately achieving very accurate use of the language. Luisa's earlier achievement of an L1 level of lexical diversity and her greater overall accuracy of language use do not appear to give her a marked advantage over the other two learners in the acquisition of new tense-aspect forms - as will be seen from the next two sections which deal with the development of past and future temporal expression.

9.3 Past-Related Morphology

Four of the research questions chosen for this study concern the development of the learners' past-related temporal expression. This section addresses these questions, discussing first the emergence of the Present Perfect (9.3.1) and then the spread of past-related forms from prototypical to less typical uses (9.3.2). Finally, the findings are summarized in Section 9.3.3.

9.3.1 Emergence of the Present Perfect

As discussed in Section 3.5, Bardovi-Harlig (2000) reports that the Present Perfect emerges after the Simple Past and Past Progressive in the

L2 acquisition of English, and that its emergence depends on the stability of the Simple Past.

Between them, Luisa, Sara and Ana create 3750 contexts for the Simple Past in their written data and as a result there is a substantial amount of information available on the learners' use of this form. Initially, Luisa's rate of appropriate use is 87% and after that it never falls below 90% (Table 7.1, Figure 7.2). Sara's initial rate is 89% and from then on is almost always over 80%. Given that 80-90% appropriate use is often considered to indicate acquisition (Bardovi-Harlig, 2000; Brown, 1973; Ellis and Barkhuizen, 2005) this would suggest that Simple Past is well established in these two learners' tense-aspect systems at the start of the study. Initially Ana's use of the Simple Past is not very stable. In Period 1 her appropriate use rate is 75%, dropping to 64% in Period 2. However from the middle of the first year of the study (Period 3 on) her appropriate use rate rarely drops below 88%.

So the Simple Past is stable in the interlanguage of all three learners at the time the Present Perfect emerges - at the end of the first year for Luisa, early on in the second year for Sara and at the end of the second year for Ana. The data confirms Bardovi-Harlig's findings in this respect. Not only is the Simple Past stable at this point, however, but it has been so for around a year or more in the case of all three learners. This suggests that while the stability of the Simple Past may be a necessary prerequisite, the emergence of the Present Perfect does not necessarily follow soon after the achievement of high rates of appropriate use of the Simple Past. This could suggest that learners need a lot of exposure to the form in order to acquire its complex meaning associations, a process made more difficult by the fact that the Present Perfect is relatively infrequent in the input, many times less frequent than the Simple Past (see below for a discussion).

Both the Simple Past and Past Progressive emerged in the learners' interlanguage before the study commenced, therefore this data cannot confirm that the simple form emerges before the progressive one. However, the data does show that the Present Perfect emerges after both past forms, as reported by Bardovi-Harlig, and that in the case of these three learners, not only the Simple Past, but also the Past Progressive is well established, with very high rates of appropriate and accurate use, before the Present Perfect emerges (and in the case of Luisa and Ana, before the Present Perfect's first recorded use in the data).

Morphosyntactic complexity and semantic complexity have both been considered as possible explanations for the later emergence of the Present Perfect (Bardovi-Harlig, 2000). From the data in this study it seems unlikely that morphosyntax is a significant factor. As already noted, the learners have been using the progressive's *auxiliary+ participle* construction accurately right from the start of the study, and there is little evidence of them struggling with the Present Perfect's *auxiliary +participle* construction - in fact it is accurately formed 93% of the time, the same percentage found in Bardovi-Harlig's study.

However semantic complexity is an issue for the learners as they do clearly struggle initially with the form-meaning associations of the Present Perfect - with identifying the different ways in which it links past events to the present moment, and with isolating when the form is needed to mark present relevance and when it is not (see 7.4.3 for examples). That the learners are still in the process of acquiring these form-meaning associations after four years exposure to English (five in Sara's case) can be seen from the fact that all three learners fail to supply the form in some of the Present Perfect contexts they create in the final month of the study. This suggests that semantic complexity may be an important factor in determining the order of emergence of this form, especially when combined with frequency of input - the form being much

less frequently used in L1 English than the Simple Past is for example (Biber et al., 1999) .

9.3.2 The spread of forms from typical to less typical uses and associations.

As a learner's tense-aspect system develops, so does their ability to express themselves with greater clarity and precision. Sometimes this expansion of communicative possibilities occurs because a new form emerges, and sometimes it occurs because an existing form takes on a new role or acquires new associations (Leech, 2004) Looking at the way a form spreads from prototypical to less typical environments can contribute to a fuller understanding of how a learner's language develops towards the norms of the target language (see 3.5.4).

The Simple Past form

The Aspect Hypothesis (Andersen and Shirai, 1994) predicts that the first verbs to attract Simple Past marking will be achievement and accomplishment verbs whose inherent meaning implies an end point; activity verbs come next and finally state verbs . By the time the study begins Luisa, Sara and Ana are routinely using Simple Past marking with all four verb classes. The data therefore is not available to explore the potential early influences of lexical aspect on verbal morphology acquisition. However, the data does allow for an analysis of potential influences of lexical aspect on later aspects of acquisition. Ana's data suggests that the spread of the form to activity and state verbs does not depend on the stability of the Simple Past. In the first half of year one of the study, 20% of the verbs found with past tense marking in Ana's data are state verbs, and between them, state and activity verbs account for 54% of the total (7.2.3). During this time Ana's appropriate use rate is as low as 64% (Table 7.1). However an analysis of Simple Past errors does provide some evidence that lexical aspect continues to exert an influence well into the final stages of the acquisition of the past tense. During the first two years of the study, on the relatively infrequent

occasions (61/1591) when base/Simple Present and *–ing* forms are found in Simple Past contexts in the learners' data, it is most likely to be activity and state verbs which are involved. They account for 77% of all such errors (7.2.3).

Andersen and Shirai (1996) hypothesized a sequence for past tense, from prototype to most marginal member, with habitual use coming after use with state verbs and before counterfactual use . The spread of the form in the learners' data follows this sequence (see 7.2.5). Andersen and Shirai predict that the habitual use will appear early on, but develop slowly, and this is the case here with it only becoming firmly established in Luisa and Ana's interlanguages in Period 5, in the second year of the study and for Sara in Period 7. Sara is more likely than the other two learners to use alternative (but appropriate) forms like *used to* and *would +verb* in habitual past contexts, and this may account – at least in part – for the habitual Simple Past taking half a year longer to become established in her interlanguage. For Luisa and Sara the counterfactual use of the Simple Past form emerges after the establishment of the habitual use. Ana creates counterfactual contexts but does not supply the appropriate form.

The Past Progressive form

The auxiliary *be + -ing* construction most typically designates an event or state of affairs in progress or continuing, and the Aspect Hypothesis predicts that the form will be found initially with activity verbs, eventually spreading to event verbs. The hypothesis is supported by evidence of a distributional bias towards activity verbs in the interlanguage of L2 learners of English (e.g. Housen, 2002; Collins, 2002).

At the start of the study, Luisa, Sara and Ana are in the later stages of the acquisition of the Past Progressive and while there is evidence of a distributional bias, it is a small one. Activity verbs are found in just over half their Past Progressive contexts (and half their Present Progressive ones). In comparison, Huang (1999) found that just under half of all

(present and past) progressive marking in her L1 data occurred with activity verbs. It would seem that in this respect Luisa, Sara and Ana's usage is close to the norms of the L1 – and is so from the start of the study, with the ratio of activity to event verbs varying very little over the four years.

Although some state verbs can combine with progressive aspect where an activity meaning is possible, statives are the most marginal member of the category 'progressive' according to Andersen and Shirai (1996) and would be expected to emerge late in a learner's interlanguage and to be used relatively rarely. In fact they are very rare in the data. In the 246 Past Progressive contexts created by the learners, state verbs occur only twice (used appropriately both times – see 7.3.4).

Romer (2005) found that in English the progressive is used much more frequently to report repeated actions or events than had previously been thought. Her corpora-based study revealed that in L1 English the progressive occurs with a 'repeatedness' meaning about 39% of the time. The learners' data suggests that at the end of the four years they are still in the process of acquiring this form-meaning association. While it emerges early in the study for all three, it is underrepresented in their interlanguage in comparison to the L1 norm.

'Repeatedness' features in only 10% of Luisa's use of the Past Progressive, 9% of Ana's and 4% of Sara's (Sara's use is higher with the Present Progressive at 10%, Ana's and Luisa's about the same for both). Often when they create a potential context for the Past Progressive with habitual or iterative meaning they supply the Simple Past or *would* +verb or *used to* + verb or lexical means instead (see 7.3.5 for examples) , and while the option chosen is sometimes inappropriate, it is generally acceptable, if not always apposite.

It is possible that, because they have these other means of expressing 'repeatedness' at their disposal, the learners need to be exposed to a considerable amount of L1 input in order to fully acquire the L1 speaker's contrastive use of the 'repeatedness' meanings of the progressive and the other forms.

The Present Perfect form

According to Leech (2004) there are two distinct ways in which a past event may be related to the present by means of the Present Perfect - it may involve a time period leading up to the present and it may have results persisting at the present time (7.4.1) . Of the four different uses of the form, 'resultative past' is the most frequently occurring in L1 English, followed by 'indefinite past'. Luisa, Sara and Ana employ the form with both of these meanings once the Present Perfect emerges in their interlanguages, and the relative frequencies found in L1 are also reflected in their data. 'resultative past' accounts for 56% of all use and ' indefinite past' for 35%.

The form can also report a state up to the present, or a habit up to the present. These two associations are 'considerably less frequent' in L1 English, according to Leech (2004) and this is also the case in the learners' data where, taken together, they amount to only 9% of all use of the Present Perfect. These uses also occur late - first appearing in the middle of the third year of the study in Luisa's case, and the middle of the fourth year in Sara's.

The Present Perfect itself emerges late for Ana, at the end of the second year, and she creates very few contexts for the form in her data – only 14, compared to 34 for Luisa and 33 for Sara. Only one of these contexts is for the habitual meaning. It occurs at the very end of the study and she supplies the Simple Present instead. Ana does develop a more targeted use of the form however. Initially she is three times more likely to use it in a Simple Past context than a Present Perfect one, but by the final year

there is much less of the overgeneralization which characterized her early use.

The learners' use of adverbials with the Present Perfect reflects L1 patterns. Leech notes that with 'indefinite past' uses, the indefinite meaning is often reinforced adverbially, especially by *ever, never* or *before*. In 70% of the learners' 'indefinite past' contexts an adverb is used – almost invariably either *ever* or *never* (7.4.3). Adverbials of duration are usually required in 'state' and 'habit' contexts and here too the learners supply them. The 'resultative past' meaning needs no support from adverbials, Leech says, and generally the learners do not use them – there are just a few instances where *already* and *just* are used appropriately.

Adverbials of duration and frequency appear early in interlanguage (Klein 1995b) and have a high salience , and in the early stages of language development, the acquisition of forms like the Simple Past may be delayed as learners may pay particular attention to these high salience adverbials, ignoring the less salient tense markings (Ellis, 2008b). However it is perhaps possible that at this later stage of morphological development, the frequent co-occurrence in the input of particular adverbs and particular Present Perfect uses (as well as the non-occurrence of adverbials with the resultative past) could aid the learners in acquiring the different meanings and uses of the form.

In summary, for Luisa and Sara, all the form-meaning associations not acquired by the start of the study become established or show emergent use in the course of the study, with several new uses emerging towards the end, suggesting that their interlanguage is continuing to develop. Ana was at an earlier stage of language development than the other two at the start of the study, and while she makes a lot of progress, her use of the Present Perfect does not include the full range of meanings.

9.3.3 Conclusion

The main points in relation to each form are as follows:

Simple Past: While the form is routinely used with activity and state verbs from the start of the study, failure to supply past tense marking tends to occur predominantly with activity and state verbs up to the end of the second year, suggesting some lingering influence of lexical aspect. From the third year on, the rare instances of errors in Simple Past contexts are as likely to involve event verbs as non-event ones. During the course of the study the habitual use of the Simple Past becomes well established in the interlanguage of all three learners, and counterfactual use emerges for Luisa and Sara.

Past Progressive: The distributional bias in favour of activity verbs which is said to characterise the earlier stages of the acquisition of the progressive is not evident in the learners' interlanguage. From the start of the study the ratio of activity to event verbs in Past Progressive contexts is close to that found in Huang's L1 corpus. However while their use of the form seems quite advanced in this sense, there is evidence of continuing development in another respect. The use of the Past Progressive with a 'repeatedness' meaning emerges for all three learners during the first year, but by the end of the study is still underused in comparison with L1 norms reported by Romer (2005). This suggests that it can take a long time for learners to establish form-meaning associations in habitual and iterative contexts.

Present Perfect: All three learners use the Present Perfect with the most typical meanings of 'resultative past' and 'indefinite past' from the time that the form emerges in their interlanguage, with 'resultative past' being the most frequently used, as it is in L1 English. The considerably less common 'state up to the present' and 'habit up to the present' meanings appear much later in Luisa and Sara's data, and not at all in Ana's. Ana

does develop a more targeted, appropriate use of the Present Perfect form towards the end, in contrast to the overgeneralization that characterized most of her early use.

Four of the research questions chosen for this study refer specifically to past-related temporal expression. The findings are as follows:

1. *Does the Present Perfect emerge in the interlanguage of all three learners?*
   Yes, the Present Perfect emerges in the interlanguage of all three learners.

2. *Does the Present Perfect emerge after the Simple Past and Past Progressive, as reported by Bardovi-Harlig (2000)?*
   In the case of all three learners, the Present Perfect emerges after the Simple Past and Past Progressive. Therefore the data here supports the order of emergence for the Present Perfect reported by Bardovi-Harlig (2000).

3. *Does the emergence of the Present Perfect depend on the stability of the Simple Past, as reported by Bardovi-Harlig (2000)?*
   The Simple Past has been stable in the interlanguage of all three learners for approximately a year before the Present Perfect emerges in their written data. This is in line with Bardovi-Harlig's (2000) finding that the emergence of the Present Perfect is dependent on the stability of the Simple Past. The data additionally suggests that while the stability of the Simple Past may be a prerequisite, the emergence of the Present Perfect does not necessarily follow soon after that stability is acquired.

4. *In the case of the Simple Past, Past Progressive and Present Perfect is there evidence of continuing development in the use of these forms – of their spread from typical to less typical meaning associations and uses?*

   In the case of all three learners there is evidence of the developing use of past-related forms, of the spread of the forms to less typical uses and associations. In fact for Luisa and Sara, all the form-meaning associations not already acquired before the start of the study become established or show emergent use in the course of the study. The 'repeatedness' use of the Past Progressive is underrepresented in all three learners' data, and this may suggest that it takes a long time to acquire the habitual meaning of the form.

9.4 Future time expression

Four of the research questions selected for this study concern the acquisition of future temporality. This section begins by discussing the order of emergence of future forms in the learners' interlanguage (9.4.1) before going on to discuss how form-meaning associations are established and then revised in the course of the development of future time expression (9.4.2). Finally, the findings are summarized in Section 9.4.3.

9.4.1 Order of emergence

The Will-Future

will +infinitive is the most frequently occurring future form in English (Leech, 2004). Research suggests that it is the first grammaticalized future form to emerge in the L2 acquisition of English (Bardovi-Harlig, 2004a) and this is the case for the three learners in this study. Luisa and Sara had probably acquired the Will-Future before the start of the study as they use it with a very high degree of accuracy right from the beginning. However Ana initially uses lexically future verbs (*want, have*

to, like) or present tense forms in future contexts , with the Will-Future not emerging in her written data until late in the first year. Once it does emerge though, it is almost always supplied where a potential context has been created, suggesting that, in her case at least, the form is quickly established once it emerges.

Until the end of the second year the Will-Future accounts for more than 80% of all future time expression (Fig. 8.1) and is typically used to express predictions, but also intentions and plans. Its use is mostly appropriate, *will* +infinitive being acceptable - if not always apposite - in many future contexts (8.7). However both Luisa and Ana overgeneralize *will* to the subordinate future where L1 speakers definitely prefer the Simple Present - Luisa overgeneralizes initially and Ana continues to do so until late in the fourth year of the study. It is possible that this overuse stems from the learners regarding *will* as a general marker of the future, as a default form to be supplied in contexts not yet associated in their interlanguage with another future form.

The data does not however support the suggestion that *will* may be regarded as a *lexical* marker by learners as suggested by Bardovi-Harlig. Bardovi-Harlig bases this suggestion on the fact that *will* emerges early, is a free morpheme and in her learner corpus 'virtually lacks variants' (2004a, 140). In Luisa, Sara and Ana's writing, *will* appears from earliest use in all variant forms, and is frequently contracted when used with pronouns – in fact *will* emerges in Period 4 for Ana and by Period 5 she is using the contracted form with pronouns (see 8.7 for a discussion).

Even after the emergence of *be going to* (Fig. 8.2) the *will* form is still dominant, involved in 60% of Luisa and Ana's future time expression, and almost half of Sara's.

Subordinate Future Simple Present

The second most frequent future construction in English is Simple Present as Subordinate Future (3.6.2). It is also the second most frequent in this

data (though much less frequent than *will*). Sara is using the subordinate future accurately from the beginning of the study. However the Simple Present does not take on this future role in Luisa's interlanguage until Period 6, the middle of the second year. In the few *if* and temporal construction contexts she creates before this she uses *will* (as already noted) or a base verb form. Ana's English language development was at an earlier stage than the other two learners at the start of the study. A subordinate future context does not appear in her data until early in the second year and there are only 14 such contexts overall. The Simple Present only takes over from *will* at the very end of the study.

Futurate Present Progressive

The next present tense form to extend its meaning to the expression of future time is the progressive. The Futurate Present Progressive (3.6.2) is the third future form to emerge in Luisa's interlanguage , the joint third in Sara's and the second to emerge in Ana's. With its specialized meaning of 'a future event anticipated by virtue of a present plan, programme or arrangement' (Leech 2004 , 61) the form is not frequently used, employed in just under 7% of the future contexts created by the learners (which is higher than the 2% recorded in Bardovi-Harlig's 2004 study). It does not emerge until quite late, in the second half of the second year - Period 7 in the case of Luisa and Ana, and Period 8 in the case of Sara – but it is used appropriately from its first appearance.

The Go-Future

be going to is the second major grammaticalized future form in English, most typically used to express the future outcome of present intention or the future outcome of present cause (Leech 2004). Bardovi-Harlig found that it emerges after *will* and that it emerges quite late (2004b). In this data emergence occurs at the end of the second year, in Period 8, for Sara, and at the start of the third year, in Period 9, for Luisa and Ana.

The *be about to* construction is similar in meaning to *be going to*, except that it suggests greater immediacy (Leech, 2004). This minor form

appears in the interlanguage of all three learners soon after the emergence of the Go-Future, but is used relatively infrequently (as is the case in L1 English).

Frequency

In Luisa and Sara's data the order of frequency of the major future forms is in line with that found in L1 English (Leech 2004):

1 Will Future
2 Subordinate Future Simple Present
3 Go-Future
4 Futurate Present Progressive

The subordinate future, emerging late, is underrepresented in Ana's data however, and comes fourth.

Discussion

As already pointed out in Section 3.1, there is very little published research dealing with the L2 acquisition of future time expression in English – the principal studies are Bardovi-Harlig 2004a and 2004b. This study confirms Bardovi-Harlig's finding that of the two main grammaticalized future forms, *will* emerges first and *be going to* emerges relatively late. Her 2004 studies did not discuss the emergence of the Futurate Present Progressive, but in this data the Futurate Present Progressive emerges before the Go-Future, or at around the same time in the case of Sara. Whether there is any significance to this, whether for example common underlying factors drive the emergence of both, is something worth exploring in future research.

Bardovi-Harlig found that the Go-Future was markedly underrepresented in her learners' future time expression with *will* forms outnumbering *be going to* forms by 5.8:1, compared to the 2.5:1 ratio found in L1 English (2004b). The Go-Future is similarly underrepresented in this data where the overall ratio is 6:1 in favour of *will* (Figure 8.4). However the ratio drops to 3.2:1 after the Go-Future emerges in the interlanguages of Luisa,

Sara and Ana. In fact after the Go Future emerges, Sara and Ana's ratios are the same, or very close to the L1 norm at 2.5:1 and 2.7:1 respectively. It is not possible to compare them, in this respect, to Bardovi-Harlig's learners as she does not provide information on the post-emergence ratio in her data.

While *be going to* is a more complex form than *will+infinitive*, this study supports Bardovi-Harlig's conclusion that morphosyntactic complexity is an unlikely explanation for the lengthy dominance of the Will-Future and the late emergence of the Go-Future. Luisa, Sara and Ana are using compositional tense-aspect forms like the progressive, as well as constructions like the infinitive of purpose, from early on in the study, and the Go-Future form is used accurately from first use.

A more likely explanation may have to do with the early emerging *will*'s utility as a marker of future time, its general acceptability in a wide range of future contexts, including all Go-Future and Futurate Present Progressive contexts. With such a useful form available, the learners may be inclined to continue to rely on it until their overall language development is such that they are motivated to use morphological means to introduce greater precision and subtlety of meaning into their future expression. For Luisa, Sara and Ana this process begins when the Futurate Present Progressive emerges to express plans (in Sara's case, concurrently with *be going to* used to express the future of present cause).

Another possible factor in the late emergence of the Go-Future might have to do with the initial 'motion' meaning attached to *going to* in the learners' interlanguage. While the metaphorical extension of *going*+infinitive of purpose to 'future intention' and from there to 'prediction' may be conventionalized in the language (Bybee, 2010; Dahl, 2000a), the learners may still need exposure to a lot of input to sort out for themselves these form-meaning associations.

9.4.2 Form-meaning associations and the development of future time expression

Prediction

The Will-Future is most typically associated with prediction (Leech 2004) and this meaning accounts for around 80% of the learners' overall use of the form. (83% for Luisa, 79% for Sara and 82% for Ana). In fact *will* is more or less the learners' sole means of expressing prediction for the first two years of the study.

The option of a second means of expressing prediction arrives with the emergence of the Go-Future, and all three learners use this form both in its typical 'future of present cause' meaning (*Look at those black clouds - it's going to rain),* and in the more general ' neutral future of prediction' sense (*No one knows if there's going to be snow this Christmas*). In fact a prediction-based meaning of the form is as likely to be used by Luisa and Sara in the initial stages as the intention-based one is (though for Sara the intention-based use is more frequent overall). For both learners the typical 'present cause' prediction meaning comes first, with the form being extended to the general prediction meaning in the later stages.

Once it emerges, Ana mainly uses the Go-Future for predictions, and although it first occurs with a 'present cause' meaning, she mostly employs the form to express the neutral future of prediction. However almost all of this 'neutral future' use occurs in just one period, where she employs *be going to* more frequently than *will* , and may be partly due to her marked tendency to temporarily overuse or make extensive use of a new form (see 8.6.2 for a discussion).

The learners in this study use the Go-Future earlier and more often for prediction than Bardovi Harlig's learners do, but as discussed in Section 8.2, prediction contexts outnumber intention ones by a ratio of 2.7:1 in this data, probably because of the wide range of writing tasks sampled, whereas in Bardovi Harlig's data - mainly drawn from journals and

interviews - the learners are much more likely to be discussing personal plans and intentions.

The Will-Future remains very strongly associated with prediction, accounting for 83% of such contexts (113/136) after the emergence of *be going to*. However the Go-Future enables the learners to link prediction to present circumstances, adding greater precision to their future time expression. They also extend their use of the Go Future to more general prediction contexts. This usage was probably present in the input the learners were exposed to, as Leech (2004) notes that the usage is increasingly found in L1 English, particularly in informal spoken English.

Plans

Until about half way through the study, the learners depend almost exclusively on *will* not only for predictions, but also to express intentions and plans. Then the Futurate Present Progressive emerges in 'plan' contexts. It is after this, or at the same time in Sara's case, that the Go Future emerges.

The 'plan' meaning is strongly associated with the Futurate Present Progressive. After it emerges, the *will* form is only found once in the data with a plan-based meaning, and *be going to* 3 times; in comparison the Futurate Present Progressive is used 28 times with this meaning. Leech notes that the difference between 'arrangement' and 'intention' is a slight one and yet in this respect the Futurate Present Progressive is almost always used very precisely by the learners, referring to a situation which has been planned, and often to a social arrangement already made (8.5.2) . The preciseness of their use of the form is interesting, given that they did not receive any explicit instruction in the Futurate Present Progressive and as a result probably acquired the form-meaning association from exposure to input alone. There are two occasions where the form is used in a context where it is referring to an intention rather than an arrangement (8.5.2) but in both cases it is expressing an intention

to *go* somewhere, and in L1 English there is a marked tendency to avoid a '*going to go*' construction by using the Futurate Present Progressive instead (Dahl, 2000a). Therefore these two instances of intention-based use of the form are target-like.

Intentions

Intention is one of the typical meanings associated with the Go-Future (Leech, 2004) and once *be going to* emerges it is almost always the form used by Ana and Sara for intention-based future expression (8.6.2). In Luisa's data *will* continues to be the main means of expressing intention, even after the Go-Future emerges. However it is always a specialized use of *will* which she employs – what Leech calls the 'performative *will*' - a use of *will* appropriate to an intention context, and particularly appropriate to the contexts she creates, introduced by verbs like *decide* and *promise* (8.3.2). On the few occasions where she is simply expressing a future intention she uses the Go-Future (8.6.2).

Intention-based reference mainly occurs in this data with first person subjects, in contrast to prediction-based reference where third person subjects predominate. This reflects a pattern found in a number of languages (Dahl, 2000a).

Imminence

The *be going to* form's most typical meanings involve a link with the present moment, with the current circumstances out of which the future arises. The future in question is often a near one, but it can be remote (Leech, 2004). In practice, most of the 'present intention' and 'present cause' uses in the data refer to the near future (today, tomorrow). And even when the learners use the Go-Future for the ' neutral future of prediction' it is mostly in relation to a near future. In fact the feature which best characterizes all three learners' use of the form is 'imminence'. It is associated with 82% of all future expression involving *be going to* (41/50).

The Futurate Present Progressive is often used in English to talk about events which are planned or arranged to occur soon, but imminence is not a necessary semantic accompaniment of the form (Leech 2004) and the learners use it almost half the time to refer to more remote events.

The Will-Future can be used to express both a near and a distant future (Leech, 2004) and in this data it is associated with imminence as well as remoteness, especially during the time that it is the main form used for intentions as well as predictions. However once *be going to* emerges, there is a noticeable shift in the learners' use of *will* – it becomes more strongly associated with remoteness (Table 9.1).

Table 9.1

% of *will* form used to express a near future before and after *be going to* emerges

	BEFORE	AFTER
LUISA	38%	20%
SARA	37%	17%
ANA	26%	16%

near future = today, tomorrow

Summary

The data clearly shows *will*'s long dominance of future expression, with the form initially used to express plans and intentions as well as predictions. The data also shows how new forms emerge to express meanings specific to them, but also, in the case of the Go Future, to compete with *will* on its own territory.

When the Futurate Present Progressive emerges in this data it almost completely takes on the role of reporting plans and arrangements – the role it fulfils in L1 English. The Go-Future when it emerges, moves into both prediction and intention contexts and is used in an appropriate, target-like way in both. Sara mainly uses *be going to* for intentions, and Ana for predictions, with Luisa's smaller sample roughly divided between the two meanings. However all three learners strongly associate the Go-Future with an imminent future, which is in keeping with the fact that imminence is a very typical (though not a necessary) semantic accompaniment of the form. *Will* retains its role as the main means of expressing the neutral future of prediction, but the future concerned is now less likely to be a close or imminent one. It is possible that as the learners' English continues to develop after the period covered by this study, their use of *be going to* will be less attached to the sense of soonness. However the Go-Future seems to have entered their interlanguage closely connected to a near future meaning, which may have helped to provide a niche for the emerging form in prediction environments previously dominated by *will*.

9.4.3 Conclusion

Four of the research questions chosen for this study refer specifically to future temporal expression. The findings are as follows:

1. *What is the order of emergence of future forms in the learners' interlanguage?*

   The data confirms that the Will Future (*will* +infinitive*)* emerges first and initially dominates future time reference, with the Go-Future *(be going to)* emerging late, as reported by Bardovi-Harlig (2004b). Additionally, in this study the Futurate Present Progressive is found to emerge before the Go Future in the case of two learners and at around the same time in the case of the third learner.

2. *If the Go-Future emerges late, is there evidence in the data to suggest which factor(s) might be responsible for its late emergence?*

   This study supports Bardovi-Harlig's conclusion (2004b) that morphosyntactic complexity is unlikely to explain the late emergence of the Go-Future. However the data does not support Bardovi-Harlig's suggestion (2004a) that *will* might initially function in the learners' interlanguage as a *lexical* marker of the future giving it an advantage over *be going to.*

3. *What meanings are initially associated with emerging future forms?*

   The data confirms that, in the absence of other future forms, the Will-Future is initially used to express intentions and plans as well as predictions, as reported by Bardovi-Harlig (2004b) . This study also finds that *will* is occasionally overgeneralized to subordinate future contexts where L1 speakers would prefer the Simple Present. From emergence, the Futurate Present Progressive is almost always used by the learners in this study to express plans, the function it performs in L1 English. While Bardovi-Harlig found that, in the early stages of emergence, the use of the Go Future is restricted to intention-based uses, and is unlikely to encode predictions (2004b, 196), in this study the learners use the Go Future from the start for both intentions and predictions, often in the form's typical 'future outcome of present circumstances' sense . Imminence is the meaning most characteristically associated with all three learners' emerging use of the Go-Future.

4. *How are form-meaning associations revised following the emergence of new forms?*

   After the emergence the Go-Future, that form becomes the main means of expressing intention. While the Will Future is still used

occasionally for intentions it is generally with its special 'performative' function. In the later stages *be going to* appears in 'general prediction' contexts and *will* is less frequently used to express an imminent future.

9.5 An overview of the order of emergence of past and future form-meaning associations

Section 9.3 looked at the development of past-related temporal expression and 9.4 at the development of future expression. The present section provides an overview of the emergence of form-meaning associations in the temporal system as a whole.

At the start of the four year study, in Period 1, Luisa and Sara's temporal expression is dominated in present and past contexts by the simple tenses, and in future contexts by *will*. The use of the Simple Past form has spread to habitual and iterative contexts, and Present and Past Progressive forms are found with a 'repeatedness' meaning (these less typical meanings of the simple and progressive forms are used relatively infrequently however). In Sara's data the spread of the simple form to a future (subordinate) context occurs during Period 1, Luisa however overuses *will* in this context.

Ana was at an earlier stage of English language development than the other two learners. While Simple Present and Present Progressive are well established at the start, it is only in Period 3, in the second half of the first year of the study, that Simple Past and Past Progressive are acquired (are used appropriately over 80% of the time). The habitual/iterative Simple Past use emerges soon after, in Period 4, followed in Period 5 by the 'repeatedness' meaning of the Past Progressive. The Will-Future does not emerge until Period 4, at the end of the first year.

For all three learners the next tense-aspect form to emerge is the Present Perfect, a form which enables the learners to link past to present. It

emerges in Period 4, for Luisa, Period 5 for Sara and Period 8 for Ana. This development is followed by the emergence of two forms which allow the learners to express a future linked to the present. The Futurate Present Progressive emerges in Period 7 for Luisa, Period 8 for Sara and Period 9 for Ana. At around the same time, or shortly afterwards, the Go-Future emerges. In Luisa and Sara's data, *will+progressive infinitive* also appears around this time, and for all three learners, the *be about to* form appears very soon after the emergence of the closely related Go-Future. Finally, the Simple Present spreads to independent clauses with future meaning in Period 10 for Luisa and in Period 12 for Sara. Although she creates some contexts for it, Ana does not employ this Futurate meaning of the Present Simple, and the use of the Simple Present in subordinate future clauses emerges very late in her data, in Period 15.

In the case of all three learners more than a year elapses between the establishment of the Will-Future and the emergence of the Futurate Present Progressive. But this emergence of the Progressive use marks the start of a series of developments in future expression which occur in quick succession, around the middle of the four year study. Most of the past-related forms analysed in this study are either already established or have shown emergent use by the end of the second year, so while the learners' use of these forms continues to development in terms of accuracy, appropriacy and appositeness, the only new form-meaning associations left to emerge in the latter part of the study are the 'state' and 'habit' meanings of the Present Perfect (Period 10 for Luisa and Period 15 for Sara). The simple and progressive forms are found in all three contexts – present, past and future – and in this data their future form-meaning associations emerge after the spread of these forms to habitual and iterative meanings.

9.6 Frequency-related factors and the learners' acquisition of English

As outlined in Section 4.2, this study adopts an Emergentist framework wherein language is a complex adaptive system which emerges from usage. Simple learning mechanisms as they are exposed to language data as part of a communicatively-rich human social environment suffice to drive the emergence of complex language representations (Ellis, 2012). The cognitive organization of language is based directly on experience with language, the basic units of language being constructions, form-meaning pairs of various levels of abstraction, and various sizes (Croft, 2001; Goldberg, 2006). In L2 acquisition (as in L1 acquisition) input and interaction are regarded as central to construction learning, with frequency being a key determinant of this learning (Beckner et al., 2009). According to Ellis (2013a) when language users demonstrate sensitivity to input frequencies this entails that they must have registered their occurrence in processing. 'These frequency effects are thus compelling evidence for usage-based models of language acquisition that emphasise the role of input' (Ellis, 2013a, 93). The analysis of the learner data in the present study provides evidence of such frequency effects.

While Luisa, Sara and Ana received some formal L2 instruction, almost all of their exposure to English took place outside the ESL classroom – in mainstream subject classes, on the school corridors and playing fields, and outside school, socializing with their friends (both native speakers and fellow L2 English speakers). The influence of the different kinds of language input they were exposed to can be seen in their data, with vocabulary, for example, ranging from the very formal to the very informal (6.2.2), the presence of adjectives such as 'pissed off' and verbs such as 'instant-message' providing clear evidence of the learners acquiring constructions without the aid of formal instruction. That learners acquire lexical items directly from exposure is not surprising, however there is also evidence in the data of the learners acquiring more

complex form-meaning associations, native-like use of temporal expression for example, without the aid of formal instruction.

The learners did not receive any explicit instruction in the Futurate Present Progressive for example, and yet with just two exceptions, this form is always used in future contexts in the data with its precise meaning of an existing arrangement made for a future event. In the other two instances, the form is used to express an intention *to go* somewhere – a usage frequent enough in L1 English to be noted in pedagogical grammars (3.6.2). The Futurate Present Progressive is not only used appropriately, it is almost always used when required once it emerges in the learners' interlanguage at the end of the second year of the study. *go* is the most frequently occurring verb in L1 use of the Futurate Present Progressive (Nesselhauf and Romer, 2007) and it is also the most frequent verb found in the learners' use of the form. That the learners' acquire appropriate use of the very specific meaning of the Futurate Present Progressive suggests a high degree of sensitivity to the pattern of use in the input of a form-meaning association which occurs relatively infrequently in future expression in L1 English (ranked joint fourth in the order of frequency of future forms by Leech, 2004).

The frequencies in the learners' data reflect the frequencies in L1 English (the BNC corpus) in a number of respects – from the relative frequencies of different tense-aspect forms, to the relative frequencies of individual words, both as they are found in particular tense-aspect contexts and also in the language data overall. In past contexts, for example, the ratio of Simple Past to Past Progressive is very similar to the ratio of simple to progressive tenses in L1 English. The Simple Past accounts for 90% of the learners' past contexts, in line with corpus findings that simple verb phrases account for about 90% of all verb phrases in L1 English (Biber et al 1999). In the learners' data the progressive accounts for 6% of past-related contexts, while in L1 English the progressive occurs in between 5% and 10% of verb phrases, depending on register (Biber et al., 1999). Until

the middle of the study the Will-Future dominates the learners' future expression (Fig. 8.1). However once the Go-Future emerges (Fig. 8.2) the frequency order for the five main future forms in the learners' data is in line with that reported for L1 English by Leech (2004) – with the one exception of Ana's infrequent use of the Subordinate Future Simple Present.

Frequency effects are also evident in an analysis of the learners' use of vocabulary. For example, in each period of the four years of the study, the majority of the ten most frequent verbs in each learner's data are verbs which also feature in the top ten on the BNC frequency list (Leech et al., 2001). An analysis of verb use in particular tense- aspect contexts also reveals frequencies in the learners' data similar to those found in L1 English. For example, in the case of each learner, seven of the ten verbs most frequently found with progressive marking are also among the ten most frequent in the BNC progressive list , with six being common to all three learners. Similarly, of the ten most frequent verbs in the Simple Past category in the BNC (Wulff et al.,2009) , seven are also found among the 12 most frequent verbs used by each learner in Simple Past contexts (Table 7.2). Overall the 20 most frequent verbs in the BNC L1 English corpus account for around 70% of all verb use in the learners' data. These verbs are likely to be very frequent in the input which the learners received and usage-based theories of language learning suggest that this would have facilitated early acquisition (Beckner et al., 2009). It could be argued that their continuing frequency of occurrence, up to the end of the four year study, is likely to be due to the fact that these verbs are generally very useful in communication, many for example being polysemous. In the later stages of acquisition Luisa, Sara and Ana may employ these verbs as often as they do for the same reasons L1 speakers do.

Corpus analysis has shown that in natural language use , word (and other construction) frequencies exhibit a Zipfian distribution (with the most

frequent word occurring approximately twice as often as the second most frequent word, three times as often as the third most frequent word, etc). It has been argued that Zipfian distributional properties help to make language robustly learnable because they result in a significant decrease in input variability (Ellis and O'Donnell, 2012). Learners tend to hear the same set of high-frequency, high utility words over and over again ; consequently, these words will be learned more quickly than if each word in the language had the same probability of occurrence (Boyd and Goldberg, 2009). As discussed in the previous paragraph, the learners' output is dominated by a small number of high frequency, high utility words. There is also evidence in the data of zipfian distribution profiles. In Simple Past contexts, for example, not only do a small number of verbs account for a large percentage of all use of the past tense, but there is also a long tail of verbs which occur only once with past tense marking. Of the 227 different verbs used in Simple Past by Luisa , the ten most frequently occurring account for more than half (53%) of the 1,436 contexts in her data, while at the other end of the scale, 99 verbs (or 43% of all verbs) are used only once. Sara's pattern of use is very similar. She employed 166 verbs with past tense marking, and her ten most frequent account for 54% of all 1165 contexts, while 41% of verbs are used only once. Ana, whose Simple Past development was the least advanced at the start of the study, relies a little more on her 'top ten' verbs – they account for 61% of her 1149 contexts – but the percentage of verbs used only once, 44%, is very similar to that found for Luisa and Sara.

While frequency is an important factor in acquisition, so is salience and the form's reliability as a predictor of an interpretation (Ellis and Collins, 2009). The data provides evidence of this, with learners struggling to acquire accurate and appropriate use of words and other constructions which are frequently occurring but have low salience and no single simple form-meaning correspondence. At the end of the study, for example, the three learners are still having difficulty sorting out appropriate contexts for frequently occurring prepositions such as *on, in* and *of* (6.3.3). The

article system is also problematical for these learners, especially so perhaps as their L1s do not have article systems (Luk and Shirai, 2009). *the* is the most frequently used word in L1 English, accounting for 7% of all word occurrences (Ellis, 2013b) and yet its use is mastered relatively late, in both first and second language acquisition (O'Grady et al., 2009). The choice of *the, a* or no article depends on a complex interaction of factors including meaning, context and shared knowledge. In relation to shared knowledge, for example, Ionin et al. (2008, 573-574) observe that 'Given the subtlety of the discourse triggers related to speaker and hearer knowledge, generalizing from them is likely to be a fairly long and difficult process'. Article system errors account for 27% of all Luisa's errors, 25% of Ana's and 19% of Sara's; while Luisa and Sara ultimately achieve over 90% accuracy, and Ana 79% accuracy, Ana is still making some relatively simple errors in the final period of the four year study.

Research on prototypes and category learning suggests that basic category terms are acquired first (Rosch et al., 1976; Tomasello, 2003). In line with this, the data for all three learners provides evidence of the acquisition of many basic category lexical items already in place by the start of the study and the new words appearing in the learners' writing for the first time in the final year tend to be infrequently occurring (in the BNC corpus), more formal or technical words. As might be expected, the more advanced the learner's acquisition of English, the higher the number of less frequent word types found in her data. 160 of the verbs Luisa uses, for example, are below 500 on the BNC frequency list, compared to 114 for Sara and 61 for Ana.

Evidence that prototypicality of meaning plays an important role in construction learning is provided by the analyses of the learners' acquisition of past and future temporal expression. The prototypical *will + infinitive*, for example, emerges first in the learners' future expression, and until the end of the second year is found in 80% of future contexts. It is suggested by this researcher (8.7) that *will*'s long dominance, and the

late emergence of the Go-Future, may be due, at least in part, to the fact that *will* is a particularly strong high-utility prototype, generally acceptable in a wide range of future contexts, including all Go-Future and Futurate Present Progressive contexts. Although there may be a loss of precision, of subtlety of meaning, when *will* is used rather than another form, *will* conveys the basic meaning. After the emergence of the other future forms, the learners' use of *will* moves closer to typical L1 usage , as the learners are less dependent on *will* for the expression of intentions (a typical meaning of the Go-Future) or plans and arrangements (the specific meaning attached to the Futurate Present Progressive).

The learners' developing use of other tense-aspect forms also provides evidence of the spread from prototypical to less typical, with the learners' data approaching closer, over time, to the L1 patterns of use. This is the case with the Present Perfect, for example. According to Leech (2004) , of the four principal meanings associated with the form, 'resultative past' is the most common, followed by 'indefinite past' . 'Resultative past' is the first meaning-association to appear in the learners' data and is the most frequently occurring in the learners' data. The indefinite past use is the second to appear, and the second most frequent in the data. The 'state' and 'habit' meanings are considerably less frequent in L1 English, and make their first appearances late in the study, being used rarely by Luisa and Sara, and not at all by Ana. The learners' use of the Simple Past shows a similar development, with the less typical habitual use becoming established in the learners' data approximately a year after stable appropriate use of the Simple Past had been achieved, and the least typical, counterfactual use, emerging in Luisa and Sara's data soon afterwards.

As the examples given in this section illustrate, the data provides evidence of the learners' sensitivity to input frequencies and of the role which frequency-related factors play in the acquisition of language from usage.

> Learners do not acquire language from unstructured, unhelpful, barren experience. Instead of desert, the evidence of language usage is as rich in latent structure as is a garden. Learners' explorations of this problem-space are grounded and contextualized. There is much latent structure to scaffold development, never mind the fact that learners' observations of this space are often directed, cultivated, and co-constructed in discourse interaction by their interlocutor as a helpful guide.
>
> (Ellis et al., 2013, 29)

10 Conclusions

10.1 Introduction

This study investigates aspects of the English language development of three teenage learners, following their progress over four school years. In this concluding chapter, Section 10.2 outlines the main contributions made by the study in relation to the findings of existing research into L2 English temporality and new findings arising from this study, a longitudinal analysis of the development of lexical diversity using vocd's D, and an analysis of the learners' orientation to the writing tasks. Some implications of the research are also briefly stated. Section 10.3 discusses proposals for further research in the areas of temporality as well as lexical diversity and suggests that the data provides sufficient evidence of sensitivity to input frequency to make a more detailed analysis of frequency effects worthwhile. The chapter ends with some concluding reflections (10.4).

10.2 Contributions to SLA Research

This study makes an important contribution to existing research on the second language acquisition of English and is, in a number of respects, innovative:

- **a four year longitudinal study**
  While the need for longitudinal research into second language development is often acknowledged in the literature, longitudinal studies are still relatively rare in this area (Dornyei, 2007). This study follows three teenage learners over a period of four school years. The four year span of the study enables it to capture the long-term nature of many developmental processes. A shorter time span might not, for example, have revealed how gradual the

acquisition of the form-meaning associations of the Present Perfect is, or how long *will* dominates the learners' future time expression.

- **a rich L2 learning corpus to add to an international database**
  The written texts, transcribed and comprehensively coded, provide a corpus of over 72,000 words. The corpus will be added to the CHILDES database (childes.psy.cmu.edu) the central repository for first and second language learning data, where it will be available to other researchers. The learners in this study come from L1 backgrounds - Malayalam, Lithuanian and Polish – which are not often found in SLA longitudinal research and this may make the corpus of particular interest to researchers investigating L1 influence on L2 acquisition.

- **an Irish post-primary school context**
  The last two decades have seen a huge increase in migration into Ireland, with many thousands of young people from a wide range of L1 backgrounds enrolled in the Irish education system. There has been relatively little published research into the L2 acquisition of English by school students in Ireland, and much that there is involves primary school pupils. This study follows the progress made by students in the language support programme of a post-primary school, where the curriculum can pose considerable challenges for L2 students.

- **a comprehensive analysis of the emergence of all future forms**
  In relation to the development of future temporality , this study confirms Bardovi- Harlig's findings that *will* emerges early and dominates future time expression, while *be going to* emerges late. This study additionally tracks the emergence of the other future forms and finds that the Futurate Present Progressive emerges before *be going to* in the case of two learners and at around the

same time in the case of the third. The study also finds that for these three learners early use of *be going to* is characteristically associated with imminence, rather than with the intention-based meaning which Bardovi-Harlig found to be characteristic of the early use of the form by most of her learners.

- **an analysis of the spread from prototypical to less typical uses**
  The investigation of the learners' past-related temporal expression includes an analysis of the spread of forms from typical to less typical associations and uses, providing an insight into the way developments continue to occur after the achievement of high rates of appropriate and accurate use of a form.

- **a longitudinal analysis of the development of lexical diversity**
  This study measures the developing lexical diversity of teenage L2 learners of English (D) providing longitudinal data over four years

- **development of a writing process designed to encourage spontaneous language use**
  The written data used in the study was drawn from a writing process developed by this researcher. The process was designed to encourage learners to focus on meaning, rather than form, in the belief that such a focus was more likely to result in the deployment of implicit linguistic knowledge (Ellis, R., 2009). This researcher then undertook an analysis of all the changes and self-corrections made by the learners in the course of writing, in order to obtain an indication of the learners' orientation to the writing task. The analysis strongly suggests that the learners were focused mainly on meaning, concerned to a much lesser extent with accuracy of form.

- **findings with implications for education policy & teaching practice**

  Luisa continued to study through Polish as well as English for the four years covered by the study, developing strong literacy skills in her L1. In contrast, all of Sara and Ana's formal education was taking place through their L2. The quicker progress made by Luisa in many areas of English language development provides some support for the Linguistic Interdependence Hypothesis proposed by Cummins (2000). The main implication of this is that schools should encourage and support L2 students in maintaining and developing literacy skills in their L1, not only for social and cultural reasons, but also to help them progress in English and achieve the best possible educational outcomes.

  In a more general way, the findings have implications both for government policy and for teaching and learning practice. The study confirms that acquisition is a slow and gradual process, much more gradual than is currently acknowledged by the Department of Education in its language support resource allocations or its pedagogical guidelines to schools and teachers.

10.3 Proposals for Further Research

Expansion of the corpus

Longitudinal studies have many strengths but they typically involve only a small number of learners, and in this study there are just three learners. However, writing samples have been collected over a four year period from two other learners in the same group as Luisa, Sara and Ana. Samples have also been collected over a two year period from four learners who arrived in Ireland at the age of 16 or 17 and enrolled in 5th year. At present samples are being collected from four further learners, with more frequent sampling periods than used before. It is proposed to transcribe and code all of this data, produced by students from a number of different L1s including Russian, Ukrainian, Kurdish, Somali, Arabic

and Amharic as well as Polish, Lithuanian and Malayalam. These transcripts would also in time be added to the CHILDES database.

Further analysis of future time expression

As the writing tasks used in this study were not designed to elicit any particular reference time or tense-aspect form, , there are less than 500 future contexts in the data. As well as analyzing the additional longitudinal data already mentioned above, it is proposed to undertake cross-sectional research specifically targeted at future time expression. Further investigation of the development of the Futurate Present Progressive is necessary to establish if the form's emergence generally occurs before or around the same time as *be going to,* as is the case for the three learners in this study. If this order of emergence is found in other learners' data then further research and analysis might be able to establish, for example, if common underlying factors drive the emergence of both. It would also be interesting to investigate further the characteristic meanings associated with early uses of *be going to*, to see , for example, if the 'neutral future' meaning which is becoming more common in spoken English (Leech, 2004) emerges as an early use in learners' interlanguage. In relation to *be going to* it is proposed that the cross-sectional research would look at patterns of use among L1 English teenage students as well as their L2 classmates.

Further analysis of the development of past-related temporality

While this study includes an analysis of the Simple Past, Past Progressive and Present Perfect, it does not consider the Past Perfect which Bardovi-Harlig reports as emerging after the Present Perfect. To complete the picture of the development of the past-related temporal expression of these three learners, an analysis of the emergence and developing use of the Past Perfect could be undertaken.

Exploration of the impact of L1 literacy skills on the development of lexical diversity in L2 English

There are a number of points of interest arising from the lexical diversity analysis in this study. For example, Luisa's data suggests that teenage learners with strong – and still developing - literacy skills in their L1 may have an advantage over other teenage learners when it comes to the development of lexical proficiency in English - and maybe even overall English language development as suggested by Cummins (2000). This is just one issue which could be examined in an expanded lexical diversity analysis, both longitudinal and cross-sectional, involving a much larger group of L2 learners, and ideally including data from L1 teenage students as well.

Frequency-related factors in language acquisition

The analysis of the data provides some evidence of sensitivity to input frequency and frequency-related factors, evidence which Ellis (2013a) argues would entail that language users must have registered patterns of occurrence in processing and so lend support to emergentist accounts of acquisition which emphasise the role of input. This is an area worth exploring further in a more focused analysis of Luisa, Sara and Ana's data. The analysis would ideally be extended to include the data already collected from the other learners mentioned above .

10.4 Concluding Reflections

I embarked on this study with two principal aims – to increase my own understanding of second language acquisition processes and to contribute to the body of knowledge in this area. I have certainly learnt a lot in the course of undertaking this research and I have been left with much to reflect on. I believe that this study does have a contribution to make to the sum of our knowledge about acquisition processes and I hope that the data collected and analysed here will form the basis for further contributions to the field.

Before I began this study I had formed two strong impressions, based on my experience as a language teacher. The first was how individual each learner is, and the second was how frequently the same patterns emerge in their learning of the language. These impressions have been confirmed for me in the course of completing this study. There is much in the data to suggest that the same underlying processes are at work in the learners' acquisition of English, but there is also evidence of individual variation – and not just in relation to orders of emergence or rates of progress, but in terms of how Luisa, Sara and Ana use the language to express themselves. Reading and rereading the transcripts I was struck many times by how each learner's use of the language is shaped by her own personality, her past experiences, her individual outlook on the world.

One final reflection - as a researcher I have an academic interest in the linguistic progress of my participants. However as a language support teacher I am very conscious of the fact that for my students, as for many thousands of immigrants in Ireland today, and many millions of people around the world, learning a second language is not just an academic exercise, but an essential part of surviving, and hopefully even thriving, in a new community. Luisa, Sara and Ana, are all well on their way.... completing their degree courses at Irish third level institutions and looking forward to their futures here in this country.

APPENDIX 1
CHAT CONVENTIONS AND CODES USED IN THE TRANSCRIPTS

The writing samples collected for this study were transcribed into CHAT format and analysed using CLAN programs. This appendix explains the conventions and codes used in these transcripts. The transcript files are provided in Appendix 2 (which is located on the attached USB). While most of the codes are standard CHAT codes, some have been adapted or are new codes devised for this particular study (see 5.3).

The three major components of a CHAT transcript are the file headers, the main tier and the dependent tiers. There are two dependent coding tiers in these transcripts – the % mor tier (see Section 5.3.3) and the %tar tier (see Section 5.3.4).

FILE HEADERS

@Languages	tells the programs which language is being used in the transcripts. For these transcripts the code is 'en' for English.
@Participants	normally lists all of the people participating in the dialogue. In this case, there is only one participant – the writer. LUS = Luisa, SAR = Sara, ANA = Ana
@ID	used to control certain CLAN programs. The format is : language|corpus|code|age| sex|role.
@Date	indicates the date on which the sample was written.
@Situation	provides background information on the writing task.

MAIN TIER

Each sentence begins with the speaker code - *ANA, *LUS or *SAR.

xxx	a partial word was erased or is illegible
xx	a word was erased or is illegible
[/]	the preceding word or scoped phrase < > was deleted and the deletion was made *before* the succeeding word was written
[//]	a change or correction (involving deletion and/or insertion) was made *after* the phrase or sentence was completed.
[= INS]	the preceding word or scoped phrase was inserted after the phrase or sentence was completed, but it does not involve a change or correction – it adds new information.
@s:	attached to the end of a word this symbol means that the learner has used an L1 word rather than an English one.
[: *word*]	following a phonetically appropriate misspelling, the correct spelling is supplied in brackets, so that the CLAN programs can recognize the word.
+'/.	used at the end of a line, this symbol indicates that a direct quotation follows.
+'	used at the start of a line to indicate that this is direct speech.
+...	used at the end of a line to indicate that the sentence was not completed

Error codes used on the main line

o*word*	a word has been omitted – e.g. *othe, oa*, indicate that the article was required, but not supplied. *odet* indicates that a determiner was omitted and is used where several options were possible e.g. *a* or *one, the or that.*
[*]	this is a general error symbol and refers to the preceding word or scoped phrase
[* os]	missing plural
[* o's]	missing possessive
[* os3]	missing 3rd person singular s
[* +s]	use of plural on singular noun

Error codes (continued)

[* +]	overuse of article
[+s3]	overuse of 3rd person singular s
[* SYN]	syntax error involving the preceding scoped phrase

MORPHOSYNTAX TIER (%mor)

The following table describes and illustrates the word class codings used for the morphosyntax analysis on the %mor tier. The table is taken from the CHAT Manual 2010 (retrieved from childes.psy.cmu.edu/manuals)

Table 50: Word Classes for English

Class	Examples	Coding of Examples
adjective	big	adj|big
adjective, comparative	bigger, better	adj|big-CP, adj|good&CP
adjective, superlative	biggest, best	adj|big-SP, adj|good&SP
adverb	well	adv|well
adverb, ending in ly	quickly	adv|quick-ADVR
adverb, intensifying	very, rather	adv:int|very, adv:int|rather
adverb, post-qualifying	enough, indeed	adv|enough, adv|indeed
adverb, locative	here, then	adv|here, adv|then
communicator	aha	co|aha
conjunction, coordinating	and, or	conj:coord|and, conj:coord|or
conjunction, subord.	if, although	conj:subord|if, conj:subord|although
conjunction, pragmatic	but	conj:prag|but
determiner, indefinite	some, any, no	det|some, det|any, det|no
determiner, singular	a, the, this	det|a, det|this
determiner, plural	these, those	det|these, det|those
determiner, conjunctive	either, neither	det|either, det|neither
determiner, possessive	my, your, her	det:poss|my
infinitive marker	to	inf|to
noun, common	cat, coffee	n|cat, n|coffee
noun, plural	cats	n|cat-PL
noun, possessive	cat's	n|cat-POSS
noun, plural possessive	cats'	n|cat-PL-POSS
noun, proper	Mary	n:prop|Mary
noun, proper, plural	Marys	n:prop|Mary-PL

Class	Examples	Coding of Examples
noun, proper, possessive	Mary's	n:prop|Mary-POSS
noun, proper, pl. poss.	Marys'	n:prop|Mary-PL-POSS
noun, adverbial	home, west	n:adv|home, n:adv|west
number, cardinal	two	num|two
number, ordinal	second	adj|second
particle	up	ptl|up
preposition	in	prep|in
pronoun, personal	I, me, we, us, he	pro|I, pro|me, pro|we, pro|us
pronoun, reflexive	myself, ourselves	pro:refl|myself
pronoun, possessive	mine, yours, his	pro:poss|mine, pro:poss|his
pronoun, demonstrative	that, this, these	pro:dem|that
pronoun, indefinite	everybody, nothing	pro:indef|everybody
pronoun, indefinite, poss.	everybody's	pro:indef|everybody-POSS
pronoun, existential	there	pro:exist|there
quantifier	half, all	qn|half, qn|all
verb, base form	walk, run	v|walk, v|run
verb, 3rd singular present	walks, runs	v|walk-3S, v|run-3S
verb, past tense	walked, ran	v|walk-PAST, v|run&PAST
verb, present participle	walking, running	part|walk-PROG, part|run-PROG
verb, past participle	walked, run	v|walk-PAST, part|run&PERF
verb, modal auxiliary	can, could, must	v:aux|can, v:aux|could, v:aux|must

%TAR TIER

The tense-aspect contexts created by the learners in the course of their writing were identified and coded on a second dependent tier (see Section 5.3.4). For example *He walked there* would be coded $POS:PST:REG (positive past regular).

POS	positive statement (POSR = in reported speech)
NEG	negative statement (NEGR = in reported speech)
NEGP	statement with *never, nobody* : *I never go; nobody saw him*
INT	question (INTN = negative question: *Don't you want to?)*
PRE	present
PST	past
FUT	future
FIP	future in past and backshifted future

%TAR TIER (continued)

MOD	modal
SIMP	simple (used for Simple Present)
REG	regular (used for Simple Past)
IRR	irregular (used for Simple Past)
PERF	perfect
PROG	progressive (used for Past Progressive and perfect progressives)
CONT	continuous (used for Present Progressive)
COP	copula *be*
WIL	will future
BGT	be going to
PRT	futurate use of present form (e.g. Futurate Present Progressive)
NA	not analyzed (used for ambiguous contexts and for contexts not analysed for this study).
[^c]	this symbol marks the beginning/end of clauses to which the %tar coding refers

SAMPLE FROM A TRANSCRIPT

```
@Begin
@Languages:    en
@Participants:  ANA Target_Teenager
@ID:    en|mcgarry|ANA|18;|female|||Target_Teenager||
@Date:  01-FEB-2011
@Situation:    The students are generating ideas for  stories.  They
      have been given about 35 minutes to write a short piece which
      involves a journey of some kind. This may later be developed into a
      full short story.
*ANA:   +" I'm bored !
%mor:   pro|I~v|be&1S adj|bored !
%tar:   $POS:PRE:COP
*ANA:   my cousin started to moan |.
%mor:   pro:poss:det|my n|cousin v|start-PAST inf|to v|moan +".
%tar:   $POS:PST:REG
*ANA:   +" let's do something .
%mor:   v|let~pro|us v|do pro:indef|something .
%tar:   NA
*ANA:   three of us were sitting on the coach [: couch] and watching tv
      [^c] because we [/] there was nothing else to do .
%mor:   det:num|three prep|of pro|us aux|be&PAST part|sit-PROG prep|on det|the n|couch
      conj:coo|and part|watch-PROG n|tv conj:subor|because pro:exist|there v|be&PAST&13S
      pro:indef|nothing post|else inf|to v|do .
%tar:   $POS:PST:PROG  [^c] $NEGP:PST:COP
```

REFERENCES

Andersen, R. W. (1977) 'The impoverished state of cross-sectional methodology (or:the left-overs are more nourishing than the main course)'. *Working Papers on Bilingualism*, 14, 47-82.

Andersen, R. (ed.) (1984) *Second Languages: A Cross-linguistic Perspective.* Rowley, MA: Newbury House.

Andersen, R. W. (1990) 'Models, processes, principles, and strategies: Second language acquisition in and out of the classroom'. In VanPatten and Lee, 45-78.

Andersen, R. W. (1991) 'Developmental sequences: the emergence of aspect marking in second language acquisition'. In Huebner and Ferguson, 305-324.

Andersen, R. W. (2002) 'The dimensions of pastness'. In Salaberry and Shirai, 79-106.

Andersen, R.W. and Shirai,Y. (1994) 'Discourse motivations for some cognitive acquisition principles'. *Studies in Second Language Acquisition*, 16, 133–156.

Antinucci, F., and Miller, R. (1976) 'How children talk about what happened'. *Journal of Child Language,* 3, 167–189.

Asher, R. E. and Kumari, T.C. (1997) *Malayalam.* London: Routledge

Ayoun, D. (2014) 'The acquisition of future temporality by L2 French learners'. *Journal of French Language Studies,* 24:2, 181–202.

Ayoun, D. and Salaberry,M.R. (2008) 'Acquisition of English Tense-Aspect Morphology by Advanced French Instructed Learners'. *Language Learning* 58:3, 555–595

Bailey, N., Madden, C. and Krashen, S. D. (1974) 'Is there a natural sequence in adult second language learning?' *Language Learning,* 24, 235-243.

Balode, L. and Holvoet, A. (2001) 'The Lithuanian language and its dialects'. In Dahl and Koptjevskaja-Tamm, 41-80.

Bardovi-Harlig, K. (1992) 'The use of adverbials and natural order in the development of temporal expression'. *IRAL,* 30, 299–320.

Bardovi-Harlig, K. (1994) 'Reverse-order reports and the acquisition of tense: Beyond the principle of chronological order'. *Language Learning*, 44, 243-282.

Bardovi-Harlig, K. (1997) 'Another piece of the puzzle: The emergence of the Present Perfect'. *Language Learning*, 47, 375-422.

Bardovi-Harlig, K. (1998) Narrative structure and lexical aspect: Conspiring factors in second language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition*, 20, 471-508.

Bardovi-Harlig, K. (2000) 'Tense and aspect in second language acquisition: Form, meaning, and use'. *Language Learning* 50 [Supplement 1].

Bardovi-Harlig, K. (2002) 'A New Starting Point? Investigating formulaic use and input in future expression'. *Studies in Second Language Acquisition*, 24:2, 189–198.

Bardovi-Harlig, K. (2004a) 'The emergence of grammaticalized future expression in longitudinal production data'. In Overstreet et al., 115-137.

Bardovi-Harlig K. (2004b) 'Monopolizing the future: How the go-future breaks into will's territory and what it tells us about SLA'. *EuroSLA Yearbook* 4, 177-201. Amsterdam: John Benjamins.

Bardovi-Harlig, K. (2005) 'The Future of Desire: Lexical Futures and Modality in L2 English Future Expression'. In Dekydtspotter, and Sprouse, 1-12.

Bardovi-Harlig, K, (2013) 'Research Design: From text to task'. In Salaberry and Comajoan, 219-269.

Bardovi-Harlig, K. and Reynolds, D. W. (1995) 'The role of lexical aspect in the acquisition of tense and aspect'. *TESOL Quarterly*, 29, 107-131.

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D. and Schoenemann, T. (2009), 'Language Is a Complex Adaptive System: Position Paper'. *Language Learning,* 59: S1, 1–26.

Behrens, H. (ed.) (2008) *Corpora in Language Acquisition Research: History, methods, perspectives.* Amsterdam: Benjamins.

Biber, D., Johansson,S., Leech, G., Conrad, S. and Finegan,E. (1999) *Longman Grammar of Spoken and Written English*. London: Longman.

Bley-Vroman (1989) 'What is the logical problem of foreign language learning?' In Gass and Schachter, 41-68.

Bloom, L. (1970) *Language Development: Form and Function in Emerging Grammars*. Cambridge, Mass: MIT Press.

Bloom, L., Lifter, K.,and Hafitz, J. (1980) 'Semantics of verbs and development of verb inflection in child language'. *Language,* 56, 386–412.

Boyd, J. K. and Goldberg, A. E. (2009) 'Input effects within a constructionist framework'. *Modern Language Journal*, 93, 418–429.

Brooks, N. (1964) *Language and language learning: theory and practice*. (2nd ed). New York: Harcourt, Brace & World.

Brown, H., Yorio, C. and Crymes, R. (eds.) (1977) *Teaching and Learning English as a Second Language: Some trends in research and practice.* Washington, DC: TESOL .

Brown, R. (1973) *A First Language: The Early Stages*. Cambridge, Mass.: Harvard University Press.

Bybee, J. (2010) *Language, Usage and Cognition.* Cambridge: Cambridge University Press.

Cancino, H., Rosansky, E., and Schumann, J. (1978) 'The acquisition of English negatives and interrogatives by native Spanish speakers' . In Hatch, 207-230.

Cancino, H. E., Rosansky, E. J., and Schumann, J. (1974) 'Testing hypotheses about second language acquisition: the copula and negative in three subjects'. *Working Papers on Bilingualism,* 3, 80-96.

Cazden, C. (1968) 'The acquisition of verb and noun inflections'. *Child Development* 39, 433-448.

Chalmers, D.J. (1995) 'Facing up to the problem of consciousness'. *Journal of Consciousness Studies*, 2, 200-219.

Chan, H.L., Finberg, J., Costello, W. and Shirai, Y. (2012) 'Lexical aspect, morphological regularity and transfer'. In Filipovic and Jaszczolt, 181-203.

Cenoz, J. and Hornberger, N. H. (eds.) (2008) *Encyclopedia of Language and Education, Second Edition, Volume 6: Knowledge about Language*. New York: Springer

Chomsky, N. (1957) *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1970) 'Remarks on nominalization.' In Jacobs and Rosenbaum, 184-221.

Chomsky, N. (1981) *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding*.Cambridge, MA: MIT Press.

Chomsky, N. (1995) *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, N. (2000) 'Minimalist inquiries: the framework'. In Martin et al., 89-155.

Chomsky, N. (2002) *On Nature and Language.* Cambridge: CUP

Collier, V. P. (1987) 'Age and rate of acquisition of second language for academic purposes'. *TESOL Quarterly*, 21, 617-641.

Collins, L. (2002) 'The role of L1 influence and lexical aspect in the acquisition of temporal morphology'. *Language Learning* , 52, 43–94.

Collins, L. (2004) 'The Particulars on Universals: A comparison of the acquisition of tense-aspect morphology among Japanese- and French-speaking learners of English'. *The Canadian Modern Language Review* 61:2, 251–74

Collins, L., Trofimovich, P., White, J., Cardoso, W. and Horst, M. (2009) 'Some input on the easy/difficult grammar question: An empirical study'. *Modern Language Journal,* 93, 336–353.

Comrie, B. (ed.) (2009) *The World's Major Languages* (2nd edition) London: Routledge.

Cook, V.J. (2003) 'The poverty-of-the-stimulus argument and structure-dependency in L2 users of English'. *International Review of Applied Linguistics* 41, 201-221.

Corder, S. P. (1971) 'Idiosyncratic dialects and error analysis'. *International Review of Applied Linguistics* 9 149-59

Croft, W. (2001) *Radical Construction Grammar: Syntactic theory in typological perspective*. Oxford: OUP

Crossley, S. A., Salsbury, T., and McNamara, D. S. (2010) 'The Development of Polysemy and Frequency Use in English Second Language Speakers'. *Language Learning,* 60:3, 573-605.

Crossley, S. A., Salsbury, T., McNamara, D. S. and Jarvis, S. (2011) 'Predicting lexical proficiency in language learner texts using computational indices'. *Language Testing*, 28:4, 561–580

Cummins, J. (1981) 'Age on arrival and immigrant second language learning in Canada. A reassessment'. *Applied Linguistics,* 2, 132-149.

Cummins, J. (2000) *Language, Power and Pedagogy: Bilingual Children in the Crossfire.* Clevedon: Multilingual Matters

Dahl, Ö. (2000a) 'The grammar of future time reference in European languages'. In Dahl, 309-328.

Dahl, Ö. (ed.) (2000b) *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter.

Dahl, Ö. and Koptjevskaja-Tamm, M. (eds.) (2001) *Circum-Baltic languages, Volume 1: Past and present.* Amsterdam: John Benjamins

Daller, M., Turlik, J. and Weir, I. (2013) 'Vocabulary acquisition and the learning curve.' In Jarvis and Daller, 185-217.

Davies, D., Criper, C. and Howatt, A.P.R. (eds.) (1984) *Interlanguage.* Edinburgh: Edinburgh University Press.

De Angelis, G. and Dewaele, J-M. (eds.) (2011a) *New Trends in Crosslinguistic Influence and Multilingualism Research.* Clevedon, UK: Multilingual Matters

De Angelis, G. and Dewaele, J-M. (2011b) 'Introduction'. In De Angelis and Dewaele, vii-xv.

de Villiers, J.G and de Villiers, P.A (1973) 'A cross-sectional study of the acquisition of grammatical morphemes in child speech'. *Journal of Psycholinguistic Research* 2, 267-278

DeKeyser, R. (2003) 'Implicit and explicit learning'. In Doughty and Long, 313- 348.

DeKeyser, R. (2014) 'The Philosophy of Science and the Social-cognitive Dichotomy in Research in Language Learning and Teaching.' In Hulstijn et al., 365-368.

Dekydtspotter, L. and Sprouse, R. (eds.) (2005) *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference,* Somerville, MA: Cascadilla Proceedings Project.

Devitt, S. (1992) *Form and function in the developing verb system of five learners of French.* Ph.D Thesis, University of Dublin.

Dietrich, R., Klein, W. and Noyau, C. (1995) (eds.) *The Acquisition of Temporality in a Second Language.* Amsterdam: John Benjamins.

Dornyei, Z. (2007) *Research Methods in Applied Linguistics.* Oxford: OUP

Doughty, C. J. and Long, M.H. (eds.) (2003) *The Handbook of Second Language Acquisition.* Oxford: Blackwell

Dulay, H. and Burt, M. (1973) 'Should we teach children syntax?' *Language Learning, 23,* 245-258.

Dulay, H. and Burt, M. (1974) 'Natural Sequences in Child Second Language Acquisition'. *Language Learning, 24:1* 37-53.

Duran, P., Malvern, D., Richards, B. and Chipere, N. (2004) 'Developmental trends in lexical diversity'. *Applied Linguistics,* 25:2 220-242

Ellis, N. C. (1998) 'Emergentism, connectionism and language learning'. *Language Learning* 48:4, 631-664.

Ellis, N.C. (2003) 'Constructions, chunking and connectionism'. In Doughty and Long, 63-103.

Ellis, N.C. (2006) 'Selective Attention and Transfer Phenomena in L2 Acquisition: Contingency, Cue Competition, Salience, Interference, Overshadowing, Blocking, and Perceptual Learning'. *Applied Linguistics,* 27:2, 164–194

Ellis, N. C. (2008a) 'Implicit and explicit knowledge about language'. In Cenoz and Hornberger, 119-132.

Ellis, N.C. (2008b) 'The Dynamics of Second Language Emergence: Cycles of Language Use, Language Change, and Language Acquisition'. *The Modern Language Journal*, 92:2, 232-249

Ellis, N.C. (2012) 'Frequency based accounts of second language acquisition'. In Gass and Mackey, 193-210.

Ellis, N. C. (2013a) 'Frequency-based grammar and the acquisition of tense-aspect in L2 learning'. In Salaberry and Comajoan, 89-118.

Ellis, N. C. (2013b) 'Second Language Acquisition'. In Trousdale and Hoffman, 365-378.

Ellis, N. C. and Collins, L. (2009) 'Input and Second Language Acquisition: The roles of frequency, form and function. Introduction to the special issue'. *The Modern Language Journal*, 93(S) , 329-335.

Ellis, N. C. and Ferreira-Junior, F. (2009) 'Construction learning as a function of frequency, frequency distribution, and function'. *Modern Language Journal*, 93(S), 370–386.

Ellis, N.C. and Larsen-Freeman, D. (2006) 'Language Emergence: Implications for Applied Linguistics'. *Applied Linguistics 27:4,* 558–589

Ellis, N. C. and O'Donnell, M. (2012) 'Statisical Construction Learning: Does a Zipfian problem space ensure robust language learning?' In Rebuschat and Williams, 265-304.

Ellis, N. C., O'Donnell, M. B. and Romer, U. (2013) 'Usage-based language: Investigating the latent structures that underpin acquisition'. *Language Learning* 63 (S1), 25-51.

Ellis, N. C., O'Donnell, M. B., and Romer, U. (2014) 'Second language verb-argument constructions are sensitive to form, function, frequency, contingency, and prototypicality'. *Linguistic Approaches to Bilingualism*, 4:4, 405-431

Ellis, R. (2008) The Study of Second Language Acquisition (2nd ed.) Oxford: OUP

Ellis, R. (2009) 'Implicit and Explicit Learning, Knowledge and Instruction'. In Ellis et al., 3-25.

Ellis, R., and Barkhuizen, G. (2005) *Analysing Learner Language*. Oxford: Oxford University Press

Ellis, R,, Loewen, S., Elder, C., Philp, J., Reinders, H. and Erlam, R. (eds.) (2009) *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. Bristol, UK: Multilingual Matters

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D. and Plunkett, K. (1998) *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA:MIT Press.

Epstein S.D., Flynn, S. and Martohardjono,G. (1996) 'Second language acquisition: Theoretical and experimental issues in contemporary research'. *Behavioral and Brain Sciences* 19:4, 677-714.

Eubank, L. (1996) 'Negation in Early German-English Interlanguage: More Valueless Features in the L2 Initial State'. *Second Language Research* 12, 73-106,

Evans, N. and Levinson, S.C. (2009) The myth of language universals: Language diversity and its importance for cognitive science'. *Behavorial and Brain Sciences* 32:5, 429-448.

Everett, D. L. (2005) 'Cultural constraints on grammar and cognition in Piraha˜: Another look at the design features of human language'. *Current Anthropology* 46:4, 621–46

Everett, D.L (2009) 'Pirahã culture and grammar: a response to some criticisms'. *Language* 85:2, 405–442.

Filipovic, L. and Jaszczolt, K.M. (eds.) (2012) *Space and Time in Languages and Cultures : Linguistic diversity*. Amsterdam: John Benjamins.

Fionda, R. (2010) *English language support for migrant students in Irish post-primary schools: an exploration of policy & practice*. PhD Thesis, University of Dublin,

Fodor, J.D. and Sakas, W. G. (2004) 'Evaluating models of parameter setting'. *Proceedings of the 28th Annual Boston University Conference on Language Development*. Retrieved from: http://www.colag.cs.hunter.cuny.edu/pub/FodorSakasBUCLD28.pdf

Fries, C.C. (1957) *The Structure of English : an introduction to the construction of English sentences.* London: Longman. First published 1952 (New York)..

Gass, S.M. and Mackey, A. (eds.) (2012) *The Routledge Handbook of Second Language Acquisition*. New York: Routledge.

Gass, S. and Schachter, J. (eds.) (1989) *Linguistic Perspectives on Second Language Acquisition.* Cambridge: Cambridge University Press.

Gass, S. and Selinker, L. (eds.) (1992) *Language Transfer in Language Learning* (revised edition). Amsterdam: John Benjamins.

Giacalone Ramat, A., & Banfi, E. (1990). 'The Acquisition of Temporality: A Second Language Perspective'. *Folia Linguistica,* 24:3–4, 405–428.

Goldberg, A. E. (2006) *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, A. E., Casenhiser, D., and Sethuraman, N. (2004) 'Learning argument structure generalizations'. *Cognitive Linguistics*, 14, 289–316.

Goldschneider, J. M. and DeKeyser, R. M. (2001) 'Explaining the "Natural Order of L2 Morpheme Acquisition" in English: A Meta-analysis of Multiple Determinants'. *Language Learning,* 51:1, 1-50.

Granger, S., Hung, J. and Petch-Tyson, S. (eds.) (2002) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching.* Amsterdam: John Benjamins.

Hakuta, K. (1976) 'A case study of a Japanese child learning English as a second language'. *Language Learning* 26:2, 321-351.

Han, Z. (2004) *Fossilization in Adult Second Language Acquisition* Clevedon, UK: Multilingual Matters.

Harris, Z. (1951) *Methods in structural linguistics* Chicago: University of Chicago Press.

Hatch, E.M. (1974) 'Second Language Learning - Universals?' *Working Papers on Bilingualism* 3, 1-17.

Hatch, E,M. (ed.) (1978) *Second Language Acquisition: A Book of Readings.* Rowley, MA: Newbury House.

Hawkins, R. and Chan, C. Y-H. (1997) 'The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis'. *Second Language Research* 13:3, 187-226 .

Hopper, P. (1987) 'Emergent Grammar'. *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society* 139-157 Retrieved from: http://elanguage.net/journals/index.php/bls/article/viewFile/2492/2459 .

Housen, A. (2002a) 'The development of tense-aspect in English as a second language and the variable influence of inherent aspect'. In Salaberry and Shirai, 155–197.

Housen, A. (2002b) 'A corpus-based study of the L2 acquisition of the English verb system'. In Granger et al., 77-116.

Huang, C. (1999) 'Tense-aspect Marking by L2 Learners of English and Native English Speakers: Inherent Lexical Aspect and Unitary vs. Repeated Situation Types'. *Issues in Applied Linguistics* 10:2 113-129.

Huebner, T. and Ferguson, C.A. (eds.) (1991) *Crosscurrents in Second Language Acquisition and Linguistic Theories* Amsterdam: John Benjamins.

Hulstijn, J. (1997) 'Second Language Acquisition Research in the Laboratory – Possibilities and Limitations'. *Studies in Second Language Acquisition* 19:2 131-143.

Hulstijn, J., and Marchena, E. (1989). 'Avoidance: grammatical or semantic causes?' *Studies in Second Language Acquisition*, 11, 241–255.

Hulstijn, J., Young, R. F. and Ortega, L. (eds.) (2014) 'Bridging The Gap: Cognitive and Social Approaches to Research in Second Language Learning and Teaching'. *Studies in Second Language Acquisition* 36:3, 361– 421.

Ionin, T., Zubizarreta, M. and Maldonado, S. (2008) 'Sources of linguistic knowledge in the second language acquisition of English articles'. *Lingua* 118, 554-576.

Jacobs, R.A. and Rosenbaum, P.S. (eds.) (1970) *Readings in English Transformational Grammar*. Waltham, MA: Ginn.

Jarvis, S. (2000) 'Methodological rigor in the study of transfer: Identifying L1 influence on the interlanguage lexicon'. *Language Learning* 50:2 245–309.

Jarvis, S. (2002) 'Short texts, best-fitting curves and new measures of lexical diversity.' *Language Testing*, 19:1, 57–84.

Jarvis, S. (2013) 'Capturing the Diversity in Lexical Diversity'. *Language Learning* 63 (S1*)* 87–106

Jarvis, S. and Daller, M. (2013) *Vocabulary Knowledge: Human ratings and automated measures*. Amsterdam: John Benjamins.

Klein, W. (1995a) 'Frame of Analysis'. In Dietrich et al., 17-30.

Klein, W. (1995b) 'The Acquisition of English'. In Dietrich et al., 31-70.

Klein, W., Dietrich, R. and Noyau,C. (1995) 'Conclusions'. In Dietrich et al., 261-280.

Klein, W. and Perdue, C. (1997) 'The Basic Variety (or: Couldn't natural languages be much simpler?)'. *Second Language Research* 13:4, 301-347.

Kleinmann, H. (1977) 'Avoidance Behavior in Adult Second Language Acquisition'. *Language Learning*, 27:1, 93–107.

Krashen, S. (1977) 'Some issues relating to the Monitor Model'. In Brown et al., 144-148.

Krashen, S., Sferlazza, V., Feldman, L., and Fathman, A. (1976) 'Adult performance on the SLOPE test: More evidence for a natural sequence in adult second language acquisition'. *Language Learning,* 26, 145-151.

Kroll, J.F. and de Groot, A.M.B. (eds.) (2005) *Handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press.

Lado, R. (1957) *Linguistics Across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor:University of Michigan Press.

Larsen-Freeman, D. (1975) 'The Acquisition of Grammatical Morphemes by Adult ESL Students'. *TESOL Quarterly,*9:4, 409-419.

Larsen-Freeman, D. (1976) 'An explanation for the morpheme acquisition order of second language learners'. *Language Learning,* 26, 125-134.

Larsen-Freeman, D. (1997) 'Chaos/complexity science and second language acquisition'. *Applied Linguistics, 18,* 141-165

Larsen-Freeman, D. and Long, M. H. (1991) *An Introduction to Second Language Acquisition Research*. London and New York: Longman.

Lee, E-J. (2001) 'Interlanguage development by two Korean speakers of English with a focus on temporality'. *Language Learning,* 51:4, 591–633.

Leech, G. (2004) *Meaning and the English Verb.* (3rd ed). Harlow, UK: Pearson Education.

Leech, G., Rayson, P., and Wilson, A. (2001) *Word Frequencies in Written and Spoken English: Based on the British National Corpus*. London: Longman.

Leech, G. and Svartvik, J. (1994) *A Communicative Grammar of English*. London: Longman

Levine, J. (1983) 'Materialism and qualia: the explanatory gap'. *Pacific Philosophical Quarterly 64* 354-361. Retrieved from: http://www.uoguelph.ca/~ abailey/Resources/levine.pdf

Lima, S.D., Corrigan, R.L. and Iverson, G.K. (eds.) (1994) *The Reality of Linguistic Rules*. Amsterdam: John Benjamins.

Little, D. (1991) *Learner Autonomy: Definitions, Issues and Problems.* Dublin: Authentik

Little, D. Devitt, S. and Singleton, D. (1989) *Learning Foreign Languages from Authentic Texts: Theory and Practice*. Dublin: Authentik

Long, M. H. (2003) 'Stabilization and fossilization in interlanguage development'. In Doughty and Long, 487-536.

Long, M.H, and Sato, C. (1984) 'Methodological Issues in Interlanguage Studies'. In Davies et al. 253-279 .

Luk, Z. and Shirai, Y (2009) 'Is the Acquisition Order of Grammatical Morphemes Impervious to L1 Knowledge? Evidence From the Acquisition of Plural-s, Articles and Possessive 's'. *Language Learning* 59:4 721-754.

MacWhinney, B (2005) 'A Unified Model of Language Acquisition'. In Kroll and de Groot, 49-67.

MacWhinney, B. (2008a) 'A Unified Model'. In Robinson and Ellis, 341-371.

MacWhinney, B. (2008b) 'Enriching Childes for Morphosyntactic Analysis' In Behrens, 165-197.

MacWhinney, B. (2010) *The Childes Project: Tools for Analyzing Talk*. Electronic Edition: Carnegie Mellon University. Retrieved from: http://childes.psy.cmu.edu

McCarthy, P. and Jarvis, S. (2007) ' vocd: A theoretical and empirical evaluation'. *Language Testing* 24:4, 459-488.

McClelland, J.L. and Rumelhart, D.E. (eds.) (1986) *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models.* Cambridge,MA:MIT Press.
Retrieved from: http://psych.stanford.edu/~jlm/papers/PDP

McGarry, D. (1995) *Learner Autonomy: The Role of Authentic Texts*. Dublin: Authentik.

McKee, G., Malvern, D. and Richards, B. (2000) 'Measuring vocabulary diversity using dedicated software'. *Literary and Linguistic Computing* 15:3, 323–37.

Malvern, D. & Richards, B. (2002) 'Investigating accommodation in language proficiency interviews using a new model of lexical diversity'. *Language Testing* 19:1, 85–104

Malvern, D.D., Richards, B.J., Chipere, N., and Durán, P. (2004) *Lexical diversity and language development*. Basingstoke: Palgrave Macmillan

Martin, R., Michaels, D. and Uriagereka, J. (eds.) (2000) *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik* Cambridge, Mass: MIT Press.

Master, P. (1987) *A cross-linguistic interlanguage analysis of the acquisition of the English article system.* Ph.D. Thesis, UCLA. Retrieved from http://www.sjsu.edu/faculty/pmaster/Master%20Dissertation.pdf

Mellow, D., Reeder, K. and Forster, E. (1996) 'Using time-series research designs to investigate the effects of instruction on SLA'. *Studies in Second Language Acquisition,* 18:3, 325–350.

Mellow, D., and Stanley, K. (2002) 'Theory Development in Applied Linguistics: Toward a Connectionist Framework for Understanding Second Language Acquisition'. *Issues in Applied Linguistics*. 13:1, 3-39

Meisel, J. M. (1987) 'Reference to past events and actions in the development of natural language acquisition'. In Pfaff, 206-224.

Meisel, J.M. (2010) 'Functional Categories (review)'. *Language* 86:4, 975-979.

Milon, J. P. (1974) 'The development of negation in English by a second language learner'. *TESOL Quarterly,* 8:2, 137-143.

Nesselhauf, N. and Römer, U. (2007) 'Lexical-grammatical patterns in spoken English: The case of the progressive with future time reference'. *International Journal of Corpus Linguistics,* 12:3, 297-333.

Norris, J. and Ortega, L. (2000) 'Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis'. *Language Learning* 50:3, 417–528.

Nunan, D. (ed.) (1987) *Applying second language acquisition research* . Adelaide, Australia: National Curriculum Resource Centre, AMEP.

Odlin, T. (1989) *Language Transfer*. Cambridge: Cambridge University Press.

O'Grady , W., Lee, M. and Kwak, H-Y. (2009) 'Emergentism and Second Language Acquisition'. In Ritchie and Bhatia, 69-86.

Ortega, L. and Iberri-Shea,G. (2005) 'Longitudinal research in second language acquisition: Recent trends and future directions'. *Annual Review of Applied Linguistics* 25, 26–45.

Overstreet,M., Rott, S., VanPatten, B. and Williams, J. (eds.) (2004) *Form and meaning in second language acquisition.* Mahwah, NJ: Erlbaum.

Parrott, M. (2010) *Grammar for English Language Teachers*. Cambridge: CUP.

Penrose, R. (1989) *The Emperor's New Mind* Oxford: OUP.

Penrose, R., and Hameroff, F. (2011) 'Consciousness in the universe: neuroscience, quantum space-time geometry and ORCH-OR theory'. *Journal of Cosmology* 14. Retrieved from http://journalofcosmology.com.

Perdue, C. (1993) *Adult Language Acquisition: Cross-Linguistic Perspectives. [Field Methods 1]*. Cambridge: CUP

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. A. and Raichle, M. E. (1988) 'Positron emission tomographic studies of the cortical anatomy of single-word processing'. *Nature* 331, 585-589 .

Pfaff, C.W. (ed.) (1987) *First and second language acquisition processes.* Boston MA: Heinle and Heinle

Pica,T. (1983) 'Adult Acquisition of English as a Second Language Under Different Conditions of Exposure' . *Language Learning* 33, 465-495.

Pienemann, M. (1998) *Language Processing and Second Language Development.* Amsterdam: John Benjamins.

Pienemann, M. and Johnston, M. (1987) 'Factors influencing the development of language proficiency'. In Nunan, 45-141.

Pinker, S. (1999) *Words and Rules: the ingredients of language*. London: Weidenfeld and Nicolson.

Pinker, S. and Prince, A. (1994) 'Regular and irregular morphology and the psychological status of rules of grammar'. In Lima et al. 321-351.

Raichle (1998) 'The neural correlates of consciousness: an analysis of cognitive skill learning'. *Philosophical Transactions of the Royal Society*. Retrieved from: http://rstb.royalsocietypublishing.org

Radden, G, and Dirven, R. (2007) *Cognitive English Grammar* Amsterdam: John Benjamins

Ravem, R. (1968) 'Language acquisition in a second language environment'. *International Review of Applied Linguistics*, 6, 165-185.

Ravem, R. (1975) 'The development of wh-questions in first and second language learners'. In Richards, 134-155.

Rebuschat, J. and Williams, J. (eds) (2012) *Statistical Learning and Language Acquisition*. Berlin: Mouton de Gruyter

Richards, J.C. (1971) 'A non-contrastive approach to error analysis.' *ELT Journal* 25, 204-219.

Richards, J.C. (ed.) (1975) *Error analysis: Perspectives on Second Language Acquisition*. London: Longman.

Ringbom, H. (1987). *The role of the first language in foreign language learning*. Clevedon, UK: Multilingual Matters.

Ritchie, W. and Bhatia, T. (eds.) (2009) The *New Handbook of Second Language Acquistion,* Bingley UK : Emerald Publishing Group

Robinson, P. (ed.) (2001) *Cognition and Second Language Instruction.* Cambridge: Cambridge University Press.

Robinson, P. and Ellis, N.C. (eds.) (2008) *Handbook of Cognitive Linguistics and Second Language Acquisition.* London: Routledge.

Rocca, S. (2002) 'Lexical aspect in child second language acquisition of temporal morphology: A bidirectional study'. In Salaberry and Shirai, 245-280.

Rohde, A. (1996) 'The aspect hypothesis and emergence of tense distinction in naturalistic L2 acquisition'. *Linguistics,* 34:5, 1115–1138

Rohde, A. (2002) 'The Aspect Hypothesis in Naturalistic L2 Acquisition: What uninflected and non-target-like verb forms in early interlanguage tell us'. In Salaberry and Shirai, 199-220.

Romer, U. (2005) *Progressives, Patterns, Pedagogy*. Amsterdam: John Benjamins

Romer, U., O'Donnell, M.B. and Ellis, N.C. (2014) 'Second Language Learner Knowledge of Verb–Argument Constructions: Effects of Language Transfer and Typology'. *Modern Language Journal*, 98:4, 952-975.

Rosansky, E. (1976) 'Methods and morphemes in second language acquistion research'. *Language Learning* 26, 409-425

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. and Boyes-Braem, P. (1976) 'Basic objects in natural categories'. *Cognitive Psychology*, 8, 382–439.

Rumelhart, D. E., and McClelland, J. (1986) 'On learning the past tense of English verbs'. In McClelland and Rumelhart, 216-271.

Sagarra, N. and Ellis, N.C. (2013) 'From Seeing Adverbs to Seeing Verbal Morphology: Language experience and adult acquisition of L2 tense'. *Studies in Second Language Acquisition,* 35:2, 261-290.

Salaberry, M.R. (2000) 'The acquisition of English past tense in an instructional setting'. *System*, 28, 135–152

Salaberry, M.R. and Comajoan, L. (eds.) (2013) *Research Design and Methodology in Studies on L2 Tense and Aspect*. Berlin: De Gruyter Mouton.

Salaberry, M. R. and Y. Shirai (eds.) (2002) *The L2 Acquisition of Tense-Aspect Morphology*. Amsterdam: John Benjamins.

Saville-Troike, M. (2006) *Introducing Second Language Acquisition*. Cambridge: Cambridge University Press.

Schachter, J. (1974) 'An Error in Error Analysis'. *Language Learning* 24:2, 205-214.

Schachter, J. and Celce-Murcia, M. (1977) 'Some Reservations Concerning Error Analysis'. *TESOL Quarterly,* 11:4, 441-452.

Schmidt, R. (2001) 'Attention'. In Robinson, 3-32.

Schumann, J. (1987) 'The expression of temporality in basilang speech'. *Studies in Second Language Acquisition,* 9:1, 21–41.

Schwartz, B. and Sprouse, R. (1996) 'L2 cognitive states and the Full Transfer/Full Access model'. *Second Language Research* 12:1, 40-72.

Selinker, L. (1972) 'Interlanguage'. *International Review of Applied Linguistics 10,* 209-31.

Selinker, L. (1984) 'The current state of IL studies: an attempted critical summary'. In Davies et al., 332-343.

Shanks, D. R. (1995) *The Psychology of Associative Learning* Cambridge: Cambridge University Press.

Shirai, Y. (1992). 'Conditions on transfer: A connectionist approach'. *Issues in Applied Linguistics,* 3:1, 91-120.

Shirai,Y. and Andersen, R.W. (1995) 'The acquisition of tense-aspect morphology'. *Language,* 71:4, 743–762.

Skinner, B. F, (1957) *Verbal Behavior.* New York : Appleton-Century Crofts.

Smyth, E., Darmody, M., McGinnity, F. and Byrne, D. (2009) *Adapting to diversity: Irish schools and newcomer students*. Dublin : ESRI.

Spada, N. and Lightbown, P.M. (2008) 'Form-Focused Instruction: Isolated or Integrated?' *TESOL Quarterly* 42:2, 181-207.

Stauble, A. M. (1984) 'A comparison of a Spanish-English and a Japanese-English second language continuum: Negation and verb morphology'. In Andersen, 323-354.

Stone, G. (2009) 'Polish'. In Comrie, 289-304.

Tavakoli, P. and Foster, P. (2011) 'Task Design and Second Language Performance: The Effect of Narrative Type on Learner Output.' *Language Learning*, 61:S1, 37–72.

Tolentino, L. C. and Tokowicz, N . (2011) 'Across Languages, Space, and Time: A review of the role of cross-language similarity in L2 (morpho)syntactic processing as revealed by fMRI and ERP methods'. *Studies in Second Language Acquisition*, 33:1, 91-125.

Tomasello, M. (1998) 'The Return of Constructions'. *Journal of Child Language* 25:2, 431-442

Tomasello, M. (2003) *Constructing a Language: A Usage-based Theory of Language Acquisition.* Cambridge, MA: Harvard University Press.

Trousdale, G. and Hoffmann, T. (eds) (2013) *Oxford Handbook of Construction Grammar.* Oxford: Oxford University Press.

Vainikka, A. and Young-Scholten, M. (1996) 'Gradual development of L2 phrase structure'. *Second Language Research* 12:1, 7-39

VanPatten, B. (1996) *Input processing and grammar instruction in second language acquisition.* Norwood, NJ: Ablex Publications

VanPatten, B. and. Lee, J.F. (eds.) (1990) *Second language acquisition—Foreign language learning.* Clevedon, UK: Multilingual Matters. (Reprinted from *IDEAL,* 1988:3,111-138)

White, L. (2003) *Second Language Acquisition and Universal Grammar.* Cambridge: Cambridge University Press

Wode, H. (1976) 'Developmental sequences in naturalistic L2 acquisition'. Working Papers on Bilingualism 11. Reprinted in Hatch 1978, 101-117.

Wode, H. (1981) *Learning a Second Language: An Integrated View of Language Acquisition.* Tubingen: Gunter Narr Verlag.

Wode, H. (1984) 'Some theoretical implications of L2 acquisition research and the grammar of interlanguages'. In Davies et al., 253-279.

Wulff, S., Ellis, N. C., Romer, U., Bardovi-Harlig, K.,and LeBlanc, C. (2009) 'The acquisition of tense–aspect: Converging evidence from corpora, cognition, and learner constructions'. *Modern Language Journal*, 93:3, 354–369.

Wunderlich, D. (2004) 'Why assume UG?' *Studies in Language* 28:3. Retrieved from: http://user.phil-fak.uni-duesseldorf.de/~wdl/Why%20UG.pdf

Yu, G. (2010) 'Lexical Diversity in Writing and Speaking Task Performances.' *Applied Linguistics*, 31:2, 236–259

Yuste, R., and Church, G.M. (2014) 'The new century of the brain'. *Scientific American* 310:3, 22-29.