UNIVERSITY OF CAMBRIDGE

FACULTY OF MODERN AND MEDIEVAL LANGUAGES AND LINGUISTICS

THEORETICAL AND APPLIED LINGUISTICS SECTION

‘The acquisition of finiteness in English by child second language learners in instructed contexts: age of onset and L1 effects’

Athina Ntalli

MURRAY EDWARDS COLLEGE

This dissertation is submitted for the degree of

Doctor of Philosophy.

June 2020

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit of 80,000 words.

Athina Ntalli
June 2020
Cambridge, UK

The acquisition of finiteness in English by child second language learners in instructed contexts: age of onset and L1 effects

Athina Ntalli

Abstract

This thesis examines the acquisition of finiteness in English by child L2 learners by investigating the impact of the age of onset and the role of the learners' L1 on their L2 acquisition. Following Meisel's hypothesis that children older than 4 will resemble adult L2 acquisition in the domain of inflectional morphology, I investigated how two groups of children of different L1s and older than 4 learn the features of tense and agreement and whether accuracy would be declining as an effect of an older age of onset.

Participants were 73 Chinese and 74 Russian learners who were aged either 9 or 12 at time of testing and had age of onset of learning English at ages 4 and 7 respectively. Children were all EFL learners recruited from EF (English First) private afternoon English language schools in Shanghai and Moscow, where children attended classes for a few hours a week. To assess children's performance, I employed two types of tasks: two elicited production tasks whose prompts involved 3SG-agreement and past tense contexts (TEGI), and a freer type of elicitation prompting stories based on a sequence of pictures (MAIN).

Data analysis demonstrated low accuracy, high numbers of omissions, asymmetries in the acquisition of morphemes, overgeneralisation of the progressive tense in 3SG-habitual contexts, and use of the periphrastic structure '*is* + *verb(x)*'. These results show that L2 children resemble aL2 acquisition supporting Meisel's hypothesis. The empirical findings are interpreted in light of two opposing views that account for the optionality in verb inflection in L2 acquisition; the Full Access to UG and the Representational Deficit approaches; as argued data are more consistent with a representational deficit account. Older children consistently outperformed younger ones; as features are inaccessible, older learners compensate by relying on their higher cognitive abilities, learning strategies and metalinguistic skills, while younger children are mostly implicit learners using more the periphrastic structure as immersed children do. The periphrastic structure appears to be a stage in L2 development of verb morphology in English which denotes the emergence of finiteness as a category being triggered semantically through interpretable features of be. This is a first stage toward activation of uninterpretable features. Finally, signs of L1 influence became more pronounced in older learners; it was the older children showing more L1 effects, a finding which is again more consistent with a representational deficit account.

Acknowledgements

I will never forget in my life the 26th of June 2016, when I received an email from Dr Theodora Alexopoulou stating my acceptance for a PhD position at the University of Cambridge. Four years later, moved, happy and proud, I have to say a farewell to this wonderful journey that my PhD was.

This journey would not have been as such without the guidance of my supervisor Dr Theodora Alexopoulou. I would thus like to express my deepest gratitude to her for her constructive advice and helpful feedback as well as for her relentless support, encouragement, and patience throughout my PhD. There are no words to describe how much I appreciated her friendship and her generosity. She stood as a role model for me both as a person and as a supervisor.

I am also deeply indebted to my other supervisor Dr Henriette Hendriks for her useful contributions, insightful suggestions, intellectual stimulation and kindness. She was there for me whenever I needed her and was always happy to discuss any issue on my work if I needed to.

I am also extremely grateful to my advisor Prof Ianthi Tsimpli with whom discussions were more than valuable. She offered her immense knowledge and experience generously and was very supportive throughout my PhD study.

This PhD would never have been possible without the financial support (studentship) by the EF Research Lab for Applied Language Learning at Cambridge. The EF Lab, as well as the Theoretical and Applied Linguistics Section of the University of Cambridge and Murray Edwards College also sponsored my attendance to several conferences, workshops and summer schools giving me the opportunity to communicate my work with fellow researchers. I thus wish to thank them very much.

Next, I would like to extend my gratitude to all the people in EF I worked with for the organization of the data collections. Of course, I owe my thanks to all the EF schools I visited in Shanghai and Moscow, and all the staff for being helpful and kind. Last but not least, I would like to express my appreciation to all the children who participated in this research as well as their parents. I hope they enjoyed the testing session as much as I did.

Many thanks go to Dr Ana Perez. First and foremost, for her friendship. Ana was also the person who introduced me to the world of mixed effects models and her help and guidance were invaluable and I am really grateful for that.

Several people in the Theoretical and Applied Linguistics Section who were also friends of mine contributed somehow to this work. I would like to thank Dr Chris Mengying Xia for translating all the material I had in Chinese; Dr Elspeth Wilson for helping me out borrowing her voice to some recordings I created; Kate Derkach for translating all the material in Russian; Carlos Balhana for discussions we had about annotation; and Dr Xiaobin Chen for introducing me to the WebAnno annotation tool and for answering any questions I had.

There are a few more people that I feel grateful to and these are my professors throughout the years. I would especially like to thank my MA supervisor Dr Olaf Koeneman, a great supervisor who inspired me get on with linguistics and Prof. Aafke Hulk who gave me the opportunity to participate in a research project as an intern during my MA studies. Perhaps these two research experiences were the stepping stones for my entrance into the PhD.

I could not leave out of this account, my fellow linguists in the Theoretical and Applied Linguistics Section who made life in the department (and outside of it) interesting and colourful. Our 'controversial lunches' as we called them with interesting debates and stimulating discussions in Selwyn College, but also all the Friday nights, the barbecues, the coffees we shared made the PhD life very enjoyable despite being stressful at the same time.

Finally, I would like to thank my family for always being by my side and who would do whatever they could to help me achieve my goals and be happy. Last but not least, I would like to thank my partner, Loukas, for everything. To him I dedicate this thesis.

Table of Contents

CHAPTER 1

Introduction

1.0 Introduction; Aims and objectives 1

1.1 Organization of the dissertation 3

CHAPTER 2

The linguistic background

2.0 The Generative approach to language acquisition 5

  2.0.1 The Minimalist Program and the architecture of language 6

  2.0.2 The role of formal features 8

  2.0.3 The clause structure 11

2.1 The generative approach in SLA 12

2.2 The functional category of finiteness 12

  2.2.1 Tense and agreement 13

  2.2.2 Aspect: Grammatical versus Lexical 13

  2.2.3 Summary 15

2.3 Finiteness in English, Chinese and Russian 16

  2.3.1 Finiteness in English 16

  2.3.2 Tense and finiteness in Chinese 18

  2.3.3 Finiteness in Russian 20

  2.3.4 Differences and similarities between the language pairs and the learning task 23

CHAPTER 3

SLA Theories

3.0 Child L2 acquisition: Definition 25

3.1 Factors that play a role in L2 acquisition 25

3.2 The age factor in cSLA 28

  3.2.1 Age effects and the Critical Period Hypothesis 28

  3.2.2 Age and Representations 30

    3.2.2.1 Representational Deficit Hypotheses 31

      3.2.2.1.1 Failed Functional Features Hypothesis 31

      3.2.2.1.2 Interpretability Hypothesis 31

3.2.2.1.3 Contextual Complexity Hypothesis 32
3.2.2.1.4 Underspecification of AspP Hypothesis 33
3.2.2.2 Full Representation approaches 35
3.2.2.2.1 Missing Surface Inflection Hypothesis 36
3.2.2.2.2 Prosodic Transfer Hypothesis 36
3.2.2.3 Feature Reassembly Hypothesis 37
3.2.2.4 A functionalist approach: The Aspect Hypothesis 38
3.3 L1 influence in cSLA 38
3.3.1 Full Access Approaches and the view of transfer at the initial stages 39
3.3.1.1 The Full Transfer/Full Access (FT/FA) model 39
3.3.1.2 Full Access/ No Full Transfer 40
3.3.2 Partial Access Approaches and their view of transfer 40
3.3.3 Summary 40

CHAPTER 4

The acquisition of inflectional morphology by L1 and L2 learners: Previous studies

4.0 Introduction 43
4.1 The acquisition of finiteness by L1 English-speaking children 43
4.2 The acquisition of finiteness in English by L2 adults 46
4.3 The acquisition of finiteness in English by L2 children 49
4.3.1 The acquisition of finiteness in English by L2 children in immersion contexts: age of onset effects 49
4.3.1.1 Summary 56
4.3.2 The acquisition of finiteness in English by L2 children in instructed settings: age of onset effects 57
4.4 Comparison of L2 learners in immersion and instructed settings 61
4.5 Interim Summary 62
4.6 Studies investigating transfer effects 63

CHAPTER 5

The Present Study: Research Questions, Hypotheses and Methodology

5.0 Introduction 67
5.1 Research Questions and Hypotheses 67
5.2 Method 68

5.2.1 Instructional setting in EF 69
5.2.1.1. EF books and introduction of grammatical phenomena 70
5.2.2 Educational context in China and Russia 70
5.2.3 Study design & rationale 72
5.2.4 Subjects 72
5.2.4.1 Chinese children 72
5.2.4.2 Russian children 73
5.2.5 Length of Exposure & Input 74
5.2.6 Proficiency 75
5.2.6.1 Word Finding Vocabulary Task 75
5.2.6.2 Class CEFR Level 78
5.2.7 Research materials 78
5.2.7.1 Test of Early Grammatical Impairment (TEGI) 79
5.2.7.2 Multilingual Assessment Instrument for Narratives (MAIN) 81
5.2.7.3 Parental questionnaires 82
5.2.8 Data collection procedure 82
5.2.9 Data transcription 84
5.3 Summary 84

CHAPTER 6

Results on finiteness marking in elicited production

6.0 Introduction 87
6.1 Data coding and scoring of the probes 87
6.1.1 Coding and scoring of the TEGI 3SG 87
6.1.2 Coding and scoring of the TEGI PAST 88
6.2 TEGI 3SG probe verbs 91
6.3 Results on TEGI 3SG 91
6.3.1 Results on TEGI 3SG for Chinese children 91
6.3.2 Results on TEGI 3SG for Russian children 96
6.3.3 Summary 101
6.4 TEGI Past tense probe verbs 101
6.5 Results on TEGI (Past regular and irregular) 102
6.5.1 Results on TEGI-past for Chinese children 102
6.5.2 Results on TEGI-past for Russian children 108

6.5.3 Summary 113
6.6 Summary of the main results of the Chapter 113

CHAPTER 7

Results on finiteness marking in narrative production

7.0 Introduction 117
7.1 Coding and annotation of the narratives 118
7.2 Annotation tool 127
7.3 Additional measures of proficiency 128
7.3.1 Narrative length: number of clauses 129
7.3.2 Lexical Diversity: Type-Token Ratio 130
7.3.3 Syntactic complexity: number of subordinate clauses 132
7.3.4 Lexical diversity and number of subordinate clauses; is there an effect of vocabulary for older Russians' higher performance? 133
7.3.5 Summary 134
7.4 Retelling versus telling 135
7.5 Results on finiteness marking in narratives 136
7.5.1 Copula be 136
7.5.2 Auxiliary be 140
7.5.3 Other auxiliaries 142
7.5.4 Main thematic verbs 145
7.5.5 Interim Summary 147
7.5.6 Individual results concerning main thematic verbs: what do they tell us about children's performance? 150
7.6 Summary of the main results of the Chapter 155

CHAPTER 8

The periphrastic structure

8.0 The pattern 'is + verb(x)' 157
8.1 Frequency and forms 157
8.2 The pattern 'be + verb(x)': Agreement 160
8.3 The pattern 'is + verb(-x)': tense and number features 161
8.4 The lexical aspect of the periphrastic structure 163
8.4.1 The lexical aspect of the periphrastic structure. Summary 174

8.5 Summary of results of the chapter 174

CHAPTER 9

Statistical modelling for finiteness marking in TEGI tasks and narrative production

9.0 Introduction on statistical modelling 177
9.1 Variables included in the analyses 177
9.2 Linear mixed-effects models 178
9.3 Statistical analyses for TEGI tasks 180
9.3.1 Binomial logistic regression for TEGI 3SG 181
9.3.2 Binomial logistic regression for TEGI Past regular 183
9.3.3 Binomial logistic regression for TEGI (irregular) Past 185
9.3.4 Binomial logistic regression for accuracy on past tense as an effect of verb regularity 186
9.3.5 Binomial logistic regression for the asymmetry in the acquisition of features 187
9.3.6 Summary of statistical results for TEGI tasks 189
9.4 Statistical analyses in narrative production 190
9.4.1 Mixed-effects binomial logistic regression for accuracy on copula be 194
9.4.2 Mixed-effects binomial logistic regression for accuracy on auxiliary be 196
9.4.3 Mixed-effects binomial logistic regression for accuracy on main thematic verbs ...197
9.4.4 Summary of statistical results for the narratives 199
9.5 Summary of results 200

CHAPTER 10

Discussion

10.0 Introduction 205
10.1 Proficiency 205
10.1.1 Methodological issues; what does it mean to compare L2 children across countries 208
10.2 L2 development for different ages of onset of learning English in EF 209
10.3 Age effects in L2 development of children with different ages of onset 210
10.4 Performance as an effect of limited input 211
10.5 The acquisition of finiteness by L2 children 213
10.5.1 Feature accessibility as a function of age of onset and L1 214
10.5.2 Morpheme effects 224

10.6 Possible explanations for the periphrastic structure ‘is + verb(x)’ structure 225
10.6.1 L1 influence 225
10.6.2 Lexical aspect...226
10.6.3 An explanation based on the framework of the Underspecification of AspP Hypothesis 227
10.6.4 Finiteness marking 227
10.6.5 Towards a new hypothesis 228
10.7 Vocabulary and grammar learning 230
10.8 Test modalities and differences in children’s performance 231
10.9 Summary 232

CHAPTER 11

Conclusions

11.0 Conclusions 235
11.1 Pedagogical implications of the study 238
11.2 Limitations of the present study & further research 239

References 243

Appendices...261

APPENDIX A

A1. Ethics Approval letter 263
A.2.1 Consent form used prior to the amendment of the Data Protection Act. 264
A.2.2 Consent form used after the amendment of the Data Protection Act. 266
A.3 Parent information sheet 268
A.4 Background questionnaire 270

APPENDIX B

B.1 Recorded stories in retelling task 272

APPENDIX C

C.1 Data coding and scoring of the probes based on the TEGI manual 273
C.2 Results of Chinese children on TEGI probes following the TEGI manual scoring method 274

APPENDIX D

D.1 Excluded verbs in narratives 273

List of Tables

Table 2.1: The lexical aspectual classes and some examples 15
Table 2.2: Vendler's (1967) classification of verbs into lexical aspectual classes 15
Table 2.3: The morphological features under investigation, their functions, and relevant examples 16
Table 2.4: Verbal morphology in English 17
Table 2.5: The aspectual system in Chinese 19
Table 2.6: Verbal morphology in Russian 21
Table 2.7: Realization of tense and agreement features in the three languages 24
Table 4.1: Results on 3SG -s and Past of L1 Chinese advanced L2 adult studies......................48
Table 4.2: Results on 3SG -s and Past tense by child L2 learners of English.........................54
Table 4.3: Immersed L2 children vs instructed L2 children with respect to morphosyntax...62
Table 4.4: Children vs adolescents/adults with respect to morphosyntax..............................63
Table 5.1: Academic hours offered for the various levels and ages by EF in China and Russia...69
Table 5.2: Information about the book series and the time the phenomena under investigation are introduced...70
Table 5.3: Study design and criteria for inclusion applied to both Chinese and Russian learners 72
Table 5.4: Chinese subjects 72
Table 5.5: Overview table of background variables extracted from the parental questionnaires of Chinese children 73
Table 5.6: Russian subjects 73
Table 5.7: Overview table of background variables extracted from the parental questionnaires of Russian children 74
Table 5.8: Mean academic hours and ranges of attendance in EF by both Chinese and Russian children 75
Table 5.9: Group average scores (standard deviations) and their ranges in brackets concerning children's scores on the Renfrew Word Finding Vocabulary Task 76
Table 5.10: Proficiency levels corresponding to the class children attended at time of testing within groups 78
T**able 6.1**: Scoring schema used for TEGI 3SG along with examples 88
Table 6.2: Scoring schema used for TEGI Past along with examples 90
Table 6.3: Chinese children's performance on TEGI 3SG in raw scores and percentages.....92

Table 6.4: Children's use of '*be + verb'* (including both grammatical '*be + verb-ing'*, and ungrammatical '*be + verb'* in the TEGI 3SG 93
Table 6.5: Item effects analysis for '*be + verb (-ing)*' forms in TEGI 3SG 93
Table 6.6: Individual results for Chinese 9-year-olds on TEGI 3SG....93
Table 6.7: Individual results for Chinese 12-year-olds on TEGI 3SG....95
Table 6.8: Russian children's performance on TEGI 3SG in raw scores and percentages....97
Table 6.9: Children's use of '*be + verb'* (including both grammatical '*be + verb-ing'*, and ungrammatical '*be + verb'* in the TEGI 3SG 97
Table 6.10: Item effects analysis for 'be + verb (-ing)' forms in TEGI 3SG....97
Table 6.11: Individual results on TEGI 3SG for Russian 9-year-old children 98
Table 6.12: Individual results on TEGI 3SG for Russian 12-year-old children 99
Table 6.13: Target verbs in TEGI Past....102
Table 6.14: Chinese children's performance on TEGI Past....103
Table 6.15: Correct suppliance of past tense markers with regular and irregular verbs and instances of overregularization 103
Table 6.16: Raw sum and percentages of accurate answers in TEGI past per (regular) item 104
Table 6.17: Children's use of 'be + verb' (including grammatical 'be + verb-ing', and ungrammatical 'be + verb' and 'be + inflected verb') in the TEGI Past 105
Table 6.18: Analysis for item effects on the production of 'be + verb(x)' forms on TEGI Past 105
Table 6.19: Individual results for Chinese 9-year-olds on TEGI Past 105
Table 6.20: Individual results for Chinese 12-year-olds on TEGI past 107
Table 6.21: Russian children's performance on TEGI Past....108
Table 6.22: Correct suppliance of regular and irregular verbs in the past condition and instances of overregularization 109
Table 6.23: Sum of accurate answers in TEGI past per (regular) item 110
Table 6.24: Children's use of 'be + verb' (including grammatical 'be + verb-ing', and ungrammatical 'be + verb' and 'be + inflected verb') in the TEGI Past 110
Table 6.25: Analysis for item effects on the production of 'be + verb(x)' forms on TEGI Past 110
Table 6.26: Individual results for 9-year-old Russian children on TEGI past....111
Table 6.27: Individual results for 12-year-old Russian kids on TEGI past....112

Table 7.1: Coding schema in narratives; data annotated for verb type, accuracy, and type of error....120
Table 7.2: All verbs; included and excluded from scoring forms....123
Table 7.3: Total number of verbs of each verb category after all exclusions....124
Table 7.4: Classification of verbs according to the lexical aspectual class and examples....124
Table 7.5: Classification of forms of '*be + verb(-x)'*....126
Table 7.6: '*be + verb(-x)'* coded for agreement and number features....127
Table 7.7: '*be + verb(-x)'* coded for tense....127
Table 7.8: Chinese children's overall accuracy on copula *be* in narratives....137
Table 7.9: Chinese children's accuracy on copula *be* in narratives....138
Table 7.10: Russian children's overall accuracy on copula *be* in narratives....138
Table 7.11: Russian children's accuracy on copula *be* in narratives....139
Table 7.12: Chinese children's overall accuracy on auxiliary *be* in narratives....140
Table 7.13: Chinese children's accuracy on auxiliary *be* in narratives....141
Table 7.14: Russian children's overall accuracy on auxiliary *be* in narratives....141
Table 7.15: Russian children's accuracy on auxiliary *be* in narratives....142
Table 7.16: Chinese children's performance on other auxiliaries in narratives....143
Table 7.17: Chinese children's accuracy with other auxiliaries in narratives....143
Table 7.18: Russian children's accuracy on other auxiliaries in narratives....144
Table 7.19: Russian children's accuracy with other auxiliaries in narratives....144
Table 7.20: Chinese children's overall use of bound morphology on main thematic verbs (MTV) in narratives....145
Table 7.21: Chinese children's use of bound morphology on main thematic verbs in narratives....146
Table 7.22: Russian children's overall use of bound morphology on main thematic verbs in narratives....146
Table 7.23: Russian children's use of bound morphology on main thematic verbs in narratives....147
Table 7.24: Chinese 9-year-old children's performance data sorted by largest number of main verbs to smallest....150
Table 7.25: Chinese 12-year-old children's performance data sorted by largest number of main verbs to smallest....151
Table 7.26: Russian 9-year-old children's performance data sorted by largest number of main verbs to smallest....152

Table 7.27: Russian 12-year-old children's performance data sorted by largest number of main verbs to smallest.....154
Table 8.1: Raw instances and percentages of periphrastic forms with *be* in both tasks for all groups across the two L1 cohorts.....158
Table 8.2: The forms the periphrastic structure appears in.....159
Table 8.3: Raw number of instances of production of agreement error in '*be*' forms of the structure '*be + verb'*.....160
Table 8.4: Raw numbers and percentages of correct and incorrect agreement between subject and the verb '*be'* in periphrastic structures in narratives.....161
Table 8.5: Raw numbers and percentages of forms of '*be'* with respect to tense marking of the periphrastic structure in TEGI past task.....162
Table 8.6: Raw numbers and percentages of different persons/forms of *'be'* as produced by each age and L1 group.....163
Table 8.7: Children's total production of verbs in the narrative task and their distribution over aspectual classes.....164
Table 8.8: Raw numbers and percentages of morphemes, bare forms and periphrastic structure instances classified by aspectual class.....165
Table 8.9: Chinese 9-year-olds' production of verbs in the narratives task and the number of their classification into aspectual class.....166
Table 8.10: Suffixes, bare forms and periphrastic occurrences by lexical classes of verbs in raw numbers and percentages.....167
Table 8.11: Chinese 12-year-olds' production of verbs in the narratives task and the number of their classification into aspectual classes.....168
Table 8.12: Number of inflected and bare forms for verbs in each aspectual class as produced by Chinese 12-year-olds.....169
Table 8.13: Russian 9-year-olds' production of verbs in the narratives task and the number of their classification into aspectual class.....170
Table 8.14: Number of inflected and bare forms for verbs in each aspectual class as produced by Russian 9-year-olds.....171
Table 8.15: Russian 12-year-olds' production of verbs in the narratives task and the number of their classification into aspectual class.....172
Table 8.16: Number of inflected and bare forms for verbs in each aspectual class as produced by Russian 12-year-olds.....173
Table 9.1: The variables, their coding, and the levels they consisted of.....177

Table 9.2: Optimal model with accuracy on 3SG-agreement as the dependent variable; significant main effects and interactions....182
Table 9.3: Optimal model with accuracy on regular past as the dependent variable; significant main effects and interactions 184
Table 9.4: Optimal model with accuracy on irregular past as the dependent variable; significant main effects and interactions 186
Table 9.5: Optimal model with accuracy on past tense as the dependent variable; significant main effects and interactions 187
Table 9.6: Optimal model with accuracy in both 3SG-agreement and regular past verbs as the dependent variable; significant main effects and interactions 189
Table 9.7: Correlational analyses between continuous predictor variables: Pearson's correlation coefficients and p-values 191
Table 9.8: Optimal model with overall verb accuracy as the dependent variable; significant main effects and interactions 192
Table 9.9: Optimal model with accuracy on copula *be* as the dependent variable; significant main effects and interactions 195
Table 9.10: Optimal model with accuracy auxiliary *be* as the dependent variable; significant main effects and interactions 197
Table 9.11: Optimal model with accuracy on main thematic verbs as the dependent variable; significant main effects and interactions....199

List of Figures

Figure 2.1: The grammatical model 8
Figure 2.2: The core spine of a clause 11
Figure 2.3: Linguistic time expressed through tense, grammatical aspect and lexical aspect ... 15
Figure 2.4: The derivation of a simple declarative sentence in English and the workings of agreement 17
Figure 2.5: Affix lowering 18
Figure 2.6: The structure of a simple declarative clause in Russian 23
Figure 3.1: The workings of the telicity feature 34
Figure 5.1: Example of a test item of the TEGI 3SG task 80
Figure 5.2: Example of a test item of the TEGI Past task 80
Figure 5.3: An example story of the narratives: ‘The dog story’ 81
Figure 7.1: Raw average number of clauses produced by the four groups and standard errors 130
Figure 7.2: Mean type-token ratio values for each group and standard errors 131
Figure 7.3: Average dependent clause ratio per group and standard errors 133
Figure 7.4: Within group-comparisons of verb form accuracy ratios in retelling and telling conditions; means and standard errors 136
Figure 7.5: Percentages of correct forms following 3SG subjects 148
Figure 7.6: Percentages of correct forms following 3PL subjects 149
Figure 8.1: Illustration of associations between morphology and lexical aspectual classes ... 166
Figure 8.2: Graphical display of the relationships between morphology and aspectual class of the verb tokens produced by Chinese 9-year-olds 168
Figure 8.3: Graphical display of the relationships between morphology and aspectual class of the verb tokens produced by Chinese 12-year-olds 170
Figure 8.4: Graphical display of the relationships between morphology and aspectual class of the verb tokens produced by Russian 9-year-olds 172
Figure 8.5 : Graphical display of the relationships between morphology and aspectual class of the verb tokens produced by Russian 12-year-olds 174
Figure 9.1: Mean accuracy on 3SG-agreement per L1 and age group 181
Figure 9.2: Mean accuracy on regular past per L1 and age group 183
Figure 9.3: The three-way interaction between L1, age, and vocabulary score 184
Figure 9.4: Mean accuracy on irregular past per L1 and age group 185

Figure 9.5: Mean accuracy on past tense (both regular and irregular verbs) per L1 and age group 186
Figure 9.6: Mean accuracy on 3SG-agreement vs regular past per L1...............................187
Figure 9.7: Mean accuracy on 3SG-agreement vs regular past per age groups...................188
Figure 9.8: Mean overall accuracy on overt finiteness marking in narrative production per age and L1 group 191
Figure 9.9: Two-way interaction of L1 and narrative length as predictors of verb form accuracy 193
Figure 9.10: Two-way interaction of age and academic hours as predictors of verb form accuracy 193
Figure 9.11: Mean accuracy on copula be per age and L1 groups......................................194
Figure 9.12: The three-way interaction between L1, age and narrative length; narrative length was only predictive for the Russian 9-year-olds...196
Figure 9.13: Mean accuracy on auxiliary be per L1 and age group.....................................196
Figure 9.14: Mean accuracy on main thematic verbs per L1 and age group 198
Figure 9.15: Two-way interaction of age and academic hours as predictors of verb form accuracy 199

List of acronyms and abbreviations

Acc	Accusative (case)
AgrP	Agreement Phrase
aL2	Adult Second Language
AP	Articulatory-Perceptual (system)
AoO	Age of Onset
BCC	Borer-Chomsky Conjecture
CEFR	Common European Framework for Reference for languages
C	Complementiser
CP	Complementiser Phrase
CH	Chinese
CCH	Contextual Complexity Hypothesis
CI	Conceptual-Intentional (system)
CPH	Critical Period Hypothesis
cL1	Child First Language
cL2	Child Second language
df	Degrees of Freedom
D	Determiner
DP	Determiner Phrase
EEG	Electroencephalography
EF	Education First/English First
EFL	English as a Foreign Language
EPP	Extended Projection Principle
ESL	English as a Second Language
FI	Full Interpretation
FRH	Feature Reassembly Hypothesis
FT/FA	Full Transfer/Full Access
GJT	Grammaticality Judgment Task
IH	Interpretability Hypothesis
IP	Inflection Phrase
LF	Logical Form
LoE	Length of Exposure
L1	First Language
L2	Second Language

MAIN	Multilingual Assessment Instrument for Narratives
MSIH	Missing Surface Inflection Hypothesis
MP	Minimalist Program
Nom	Nominative (case)
P&P	Principles and Parameters
PF	Phonetic Form
PTH	Prosodic Transfer Hypothesis
PWd	Prosodic Word
RU	Russian
SE	Standard Error
SES	Socio-economic status
SLA	Second Language Acquisition
Spec	Specifier
T	Tense
TEGI	Test of Early Grammatical Impairment
TP	Tense Phrase
TD	Typically Developing
UG	Universal Grammar
V	Verb
VP	Verb Phrase

CHAPTER 1

INTRODUCTION

1.0 Introduction; Aims and objectives

This thesis aims to contribute to the understanding of child second language acquisition (cSLA). What makes child second language (cL2) acquisition a distinct area of research? Schwartz (2004), in a paper entitled 'Why cL2 acquisition?' argued that this population is crucial for our understanding of the other acquisition types such as child first language (cL1) and adult second language (aL2), as it shares characteristics with both: cognitively, the child L2 learner is more similar to the child L1 learner, but cL2 learners resemble aL2 learners in that they both have already acquired another language, their mother tongue or L1. However, L2 children are growing so age effects are bound to be relevant. All these make L2 children have a unique profile making their investigation interesting not only for the sake of understanding the other two acquisition types but in its own right as well.

Theoretical considerations have indeed suggested that in some domains L2 children may resemble L1 children and in others L2 adults. Schwartz (2004), for instance, proposed that L2 children will resemble L2 adults in syntax and will pattern with L1 children in inflectional morphology. On the other hand, Meisel (2009) proposed that L2 children will pattern with L2 adults in the domain of inflectional morphology, leaving open the question of what we may expect in the domain of syntax.

The impact of the age of onset of acquisition is a central question in cL2, distinguishing it from both cL1 and aL2. It relates to the long-standing question of the existence of a Critical Period for language acquisition but also age effects on the general cognitive development of young learners.

In this thesis, I focus on how age of onset impacts on the acquisition of verbal morphology on cL2. There are two reasons for investigating morphology: the first reason is theory independent and is based on previous empirical studies which have shown that morphology is particularly challenging for both aL2 (e.g. Hawkins & Chan, 1997; DeKeyser, 2005; Lardiere, 2009; Slabakova, 2016) but also cL2 learners (e.g. Paradis et al. 2008; Paradis et al. 2016). The second reason is theory-driven in that parametric variation among languages lies within formal functional features and therefore the acquisition of these features is central to a generative perspective for accounting for cL2.

Specifically, I investigate the acquisition of the verbal suffixes, third person singular -*s* and past tense. These are notoriously difficult phenomena for L2 learners as discussed in many previous studies (cf. Sections 4.2-4.3 /Chapter 4). A classic example is the (lack of) acquisition of these morphemes by Patty, a Chinese adult L2 learner of English (Lardiere, 2009). There is vast literature on the acquisition of verbal morphology which further allows for comparisons between cL2, cL1 and aL2 learners (cf. Chapter 4). Finally, inflectional morphemes are crucial for our understanding how the L1 availability of these features influences their L2 acquisition.

This brings us to another crucial question in cL2, that is, the role of crosslinguistic influence, specifically the typological impact of the learners' first language on their second, or transfer. Transfer effects can manifest as instances of L1 structures erroneously transferred in the L2 (negative transfer) or as acceleration effects (positive transfer) when the two languages share the same structures/features (Paradis, 2011). These effects have been widely documented for aL2 (Jarvis & Pavlenko, 2008; Unsworth & Hulk, 2009). Transfer is of course absent from child L1 acquisition. Thus, identifying the role of transfer and how it affects cL2 is crucial for understanding how it makes cL2 different from cL1.

I approach transfer from the generative perspective of parametric variation across languages. Thus, Chinese and Russian learners of English were chosen as the subjects of investigation. The language combinations proposed here should be able to shed light on the influence of the L1 since -as we will see in the next section- there are interesting typological differences between them. Chinese is an isolating language meaning that morphology is poor to non-existent while Russian is a morphologically rich language, that is, highly inflecting. Their verbal morphological paradigms then differ with respect to English.

The population of this study constitute English as a Foreign Language (EFL) learners. Children learn English in a classroom context attending classes for a few hours a week. By studying this population, we may be able to see whether input changes the developmental route in the acquisition of morphology L2 children go through and whether this resembles the cL1 or aL2. Research has shown that instructed learners go through the same developmental stages with naturalistic learners (e.g. Ellis, 1984a, 1989; Krashen, 1985; White et al. 1992; Bardovi-Harlig 1995). Nevertheless, the many differences between instructed and naturalistic contexts in terms of input (e.g. when and how structures are taught) in its interaction with the age of young learners can potentially influence acquisition.

Our focus on the acquisition of verbal morphology in an EFL context can potentially have pedagogical implications. It is therefore important to base any pedagogical recommendations on research combining insights from naturalistic and EFL contexts. EFL child learners constitute a rather underresearched group in linguistic SLA. This thesis aims to fulfil this gap.

To empirically address the impact of the age of onset as well as the role of the crosslinguistic influence on cL2 acquisition, I focus on how children at two different ages, speaking two different source languages whilst living in two different countries learn the features of tense and agreement in English in a classroom context. Specifically, 9- and 12-year-old Chinese and Russian children with age of onset of learning English at English First (or Education First), EF schools, in Shanghai and Moscow with age of onset at 4 and 7 respectively were tested on 3SG-agreement and past tense through (oral) elicited production tasks.

1.1 Organization of the dissertation

This thesis is organised as follows: Chapter 2 presents the generative framework or the Universal Grammar approach -that this work adopts- as understanding the architecture of language and the locus of variation across languages will help us understand what the learning task in L2 acquisition is. Chapter 3 reviews the theoretical background of this study; it specifically discusses the L2 theories and hypotheses concerning age effects and feature accessibility in both adult and child learner acquisition types. Chapter 4 provides a review of empirical studies focusing on the acquisition of tense and agreement features in English by a variety of populations. Chapter 5 presents the research questions and hypotheses of this study as well as the methodology followed to address them; the study design, the participants, the tasks, and the testing procedure. Chapter 6 presents results of production tasks eliciting third person singular -*s* and past tense while Chapter 7 presents results from narrative productions focusing again on the same features. Chapter 8 considers in detail an interesting pattern attested in children's speech but not existing in the children's input, the '*is + verb(x)*' structure. Chapter 9 provides statistical analyses for results for all tasks and Chapter 10 discusses the findings altogether in light of current child L2 hypotheses. Finally, Chapter 11 wraps up this study summarising the conclusions, discussing the study's limitations, and suggesting future research directions.

CHAPTER 2

THE LINGUISTIC BACKGROUND

2.0 The Generative approach to language acquisition

The generative framework or the Universal Grammar (UG) approach has been being developed by Noam Chomsky and his followers since the late 1950s. The main aim of this theory is to understand and explain the language system as represented in the speaker's mind (Chomsky, 1986a). Chomsky (1986a) suggested that the goals of linguistic theory are to address basic questions about human language, that is, its nature, the way it is acquired and how it is put to use. With respect to these questions, UG came as an answer. What is UG though?

UG is the centrepiece of the language faculty, the 'genetic blueprint' (White, 2003a: 2) consisting of principles which are invariant across languages and parameters where languages differ. Humans are endowed with a biological 'organ' for language or in other words, an innate language faculty (Chomsky, 1965; Pinker, 1984) which constraints or shapes what languages can look like.

This approach attempted to address the logical problem of first language learners who, despite being exposed to degenerate input, all acquire their native language's grammatical system with the same degree of success at around the same time while going through the same stages during development. Children were found to be resistant to correction, and they all, irrespective of the language learned, culture, and differences in intelligence seemed to acquire the grammar of their native language rapidly, uniformly and successfully. Evidence further suggested that little input or 'positive evidence' was needed for children to acquire very subtle linguistic phenomena. These arguments -and further ones- led Noam Chomsky and other generative linguists to speculate that there is an innate template guiding children's language acquisition, UG.

Linguistic theory has evolved through the years and in light of new languages being studied, the idea of UG has undergone conceptual changes. In the Principles & Parameters (P&P) framework (Chomsky, 1981), UG was supposed to consist of invariant principles true for all languages (e.g. the Structure Dependency Principle (i.e. the fact that sentences are organised hierarchically in terms of phrases and not individual words), the Extended Projection Principle ((EPP, i.e. the fact that all sentences must have a subject)) and a number of parameters associated to those principles explaining variation across languages and 'governing whole clusters of properties' (Chomsky, 1981a). To give an example of one such parameter, the 'verb movement parameter' was supposed to be responsible for adverb placement, negation formation, question formation, etc. (White, 1990/1991). Parameters were visualised as on-and -off switches and their setting would entail the said clustering effect (Thornton & Crain, 2013).

Principles were supposed to come for free since they were universal, while parameters were language-specific and needed to be learnt under exposure to input. When acquiring a parameter many properties associated with it would fall into place. Hence, parameters were supposed to alleviate the acquisition task in first language acquisition as children were thought to have access to the parameter inventory and with exposure to input the parameter with the associated properties would be set. Both notions were revised in the Minimalist Program (Chomsky, 1995) which we will discuss in the next section.

2.0.1 The Minimalist Program and the architecture of language

In 1995, Chomsky introduced the Minimalist Program (MP), the most recent incarnation of the generative theory. According to the MP, the language faculty consists of two main components: 1. the computational system (also called 'narrow syntax'), and 2. the lexicon. Two further interface components were also proposed: the Phonetic form (PF) and the Logical Form (LF). The lexicon contains lexical (or substantive) categories such as nouns [N], verbs [V], and adjectives [A] as well as functional (or non-substantive) categories such as determiner [D], tense [T], and complementizer [C]; the former provide the lexical content or concepts while the latter provide the grammatical information (i.e. they are grammatical elements). The functional categories are also associated with a number of formal features; for example, the functional category of tense [T] may be associated with the features [agreement, past, number, person, etc.] in different languages. As such, categories have a 'positional definition' as Adger and Svenonius (2010) put it, while features are subclassified properties of a category.

The features and concepts are selected from the lexicon and are fed into the computational system, which generates the syntactic derivation. There is a small number of syntactic operations: 'Merge', 'Move' (later renamed Internal Merge) and 'Agree'. 'Merge' is the fundamental structure-building operation that combines elements together in a pair-wise fashion according to the Binarity Condition (Chomsky, 1995). 'Move' is the syntactic operation that builds the structure so as to express 'scopal and discourse-related properties'

(Chomsky, 2001:9). In early Minimalism, 'Move' was considered an imperfection and a costly operation compared to 'Merge'. However, it was then considered an instantiation of Internal Merge losing its status as a separate operation and also being considered equal to External Merge. Finally, 'Agree' is the final operation that takes place in the narrow syntax and its role is to get together heads with features that need to be valued (i.e. uninterpretable/unvalued features) with heads that have the relevant features (i.e. interpretable/valued features) so that the former are checked-off. 'Agree' -previously known as 'feature checking'- requires movement which takes place either in the Narrow Syntax, that is before Spell-out (i.e. the point when a syntactic derivation passes on to the interface components) when there are strong features or later on at LF for the weak features. These syntactic operations are universal.

Syntax transfers the derived structure to the PF and LF interfaces which are independent from each other. The PF 'translates' the syntactic derivation in terms of 'sound' while the LF 'interprets' the product generated by syntax in terms of 'meaning' representation. The two interfaces interact with the Articulatory-Perceptual system (AP) and the Conceptual-Intentional system (CI) which are different parts of our cognition.

In this system, morphology is a separate grammatical module from syntax and PF. It follows the syntactic operations, that is, it receives the syntactic output and provides the forms. For example, morphology 'sees' a determiner phrase (DP) from the syntactic derivation carrying the features [φ=3SG, Fem] and [uFin][1] and provides the form 'she' for it (Koeneman & Zeijlstra, 2017). This happens after the syntactic derivation has been completed which works on the basis of feature bundles. 'This idea that syntax operates with bundles of features, and that morphology inserts concrete forms after syntax is done, is a core tenet of Distributed Morphology (e.g., Halle & Marantz, 1993; Noyer, 1997; Marantz, 1997, Bobaljik, 2017), a very influential proposal in theoretical morphology.

There is not one-to-one relationship between syntax and morphology, that is, morphology does not always express the formal syntactic features. For example, the first person singular of the present tense in English e.g. I read-∅ has a null (or non-overt) exponent; this does not mean that the features are not present in the syntactic derivation or that there is

[1] [uFin] stands for finite. According to the MP, it is the finite feature of the T head that assigns nominative case to the Spec, T and thus we get 'she' and not 'her'. The [finite] feature is uninterpretable on the pronoun ([uFin]) because it needs to be matched with the corresponding interpretable [Fin] feature on the verb checking the features of [tense, person, number].

no feature checking (or ‘Agree’) between the verb and the subject (in minimalist terms: head, T and the Spec, T).

Figure 2.1 below shows the grammatical model, a.k.a. Y-inverted model.

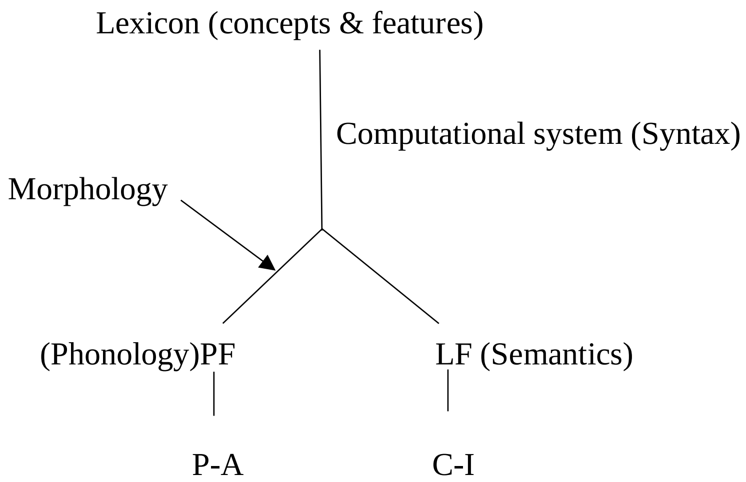


Figure 2.1: The grammatical model (slightly adapted from Koeneman & Zeijlstra, 2017)

Let us consider parametric variation across languages according to this model. Borer (1984a) first proposed that parameters are located within the lexicon, an idea adopted by Chomsky (1995) which was then labelled the ‘Borer-Chomsky Conjecture’ (BCC; as named by Baker 2008a:353). This proposal is nowadays a main assumption in the Minimalist theorising (Slabakova, 2016). Specifically, it states that the locus of parametric differences lies within the features of functional categories. In other words, each language may select a different number of features from the universal inventory to realise resulting in differences among them.

In the next section, I consider the role of features and their relation to parametric variation.

2.0.2 The role of formal features

Given the centrality of features in parametric variation, recent work in acquisition studies is essentially work on the acquisition of features as they form the basis of various L2 learnability hypotheses.

As we have seen above, syntax ignores individual words and works with (formal) features which it combines together. Features are phonological, semantic, and syntactic with only the syntactic ones being relevant to the syntactic computation. Features are thus seen as the building elements of linguistic structure (Liceras et al., 2008) or as Adger (2003) characterised them: the 'atoms' of language. A formal (or functional) feature is a property distinguishing some elements from others (Adger & Svenonius, 2010: 8). An example of a feature in English grammar is the [past] feature as manifested in the suffix *-ed*.

(1) play*ed* [V, past, perfective]

In (1) the verb form *played* is specified for [past] tense as well as [V]. The suffix *-ed*, however, is also specified for [perfective] aspect so that a single morpheme may correspond to more than one feature.

Features can be divided into two categories: intrinsic and optional. Intrinsic features are those which are stored in lexical entries (e.g. mass vs count nouns) while optional features are those which are not necessarily encoded in any occurrence of a lexical item as it is [number] on nouns. For example, the word *'worked'* has an intrinsic feature [V] and an optional feature [past], because the latter could also be marked for [3SG].

Another important distinction concerning features is based on their interpretability status distinguishing them between interpretable and uninterpretable ones (Chomsky, 1995a: 277). (LF-) Interpretable features are those which contribute to the sentence meaning as is the [plural] *-s* on nouns or the [past] feature on a verb (Adger & Svenonius, 2010). LF-interpretable features can be combined with PF-interpretable features (e.g. plural *-s)* or PF-uninterpretable features (e.g. 'animacy distinctions on Greek nouns and pronouns which are not grammaticalized due to grammatical gender differences', (Tsimpli & Dimitrakopoulou, 2007:223)). (LF-) Uninterpretable are those features which are responsible for establishing syntactic dependencies. For example, the third person singular *-s* on a verb is an uninterpretable feature because it does not contribute semantic content to the sentence meaning but is there only for syntactic purposes, that is, establishing agreement with the subject -or put it in minimalist terms, the T head and the DP in the specifier of T need to agree for the values of [person] and [number]. Other examples of uninterpretable features are the nominative [Nom] or the accusative [Acc] cases on nouns as they do not contribute with any meaning, but they are only syntactically relevant. Uninterpretable features are viewed as triggers of syntactic agreement as these features need to be matched, 'checked' or 'get valued' by the corresponding interpretable ones in the same clause and get deleted (Koeneman & Zeilstra, 2017:137; Chomsky, 2000, 2001). In other words, uninterpretable features drive the syntactic derivation. As their role is syntax-internal only, they do not survive at LF. LF-uninterpretable features may have morphophonological reflexes, so they might be PF-interpretable (3SG-agreement *-s*) or may be combined with PF-uninterpretable features (case on nouns in English) (Tsimpli &

Dimitrakopoulou, 2007). On the other hand, interpretable features are interpretable at one or more interfaces and they survive beyond the syntax at (both PF and) LF. Interpretable features are thus syntactically-relevant semantic features.[2]

Finally, features can differ in their strength being either strong or weak at least as conceptualised in early minimalism of feature-checking while in later agree-oriented minimalism strong and weak were substituted by the EPP[3] feature – a diacritic for movement. In both renditions of the minimalist programme, those features that were strong (as indicated with a star, e.g. [wh*] or otherwise said, being further specified with the EPP feature, e.g. [wh$_{EPP}$] were supposed to trigger overt movement (as in e.g. English; the wh-word needs to move to the Spec, C). Weak features or features lacking the EPP sub-feature trigger covert movement at LF after the Spell-out so that there will be no PF effects and hence in the example of the wh-movement, the wh-element would remain in situ (as in e.g. Chinese).

Let us now see how exactly variation across languages is captured as variation in features. According to White (2003),

- first, languages may differ with respect to the functional categories they realize; for example, Japanese is argued by some researchers to lack the category D (Fukui & Speas, 1986 as referenced in White, 2003) while English instantiates it,[4]
- second, languages may differ with respect to the features they realize; for instance, Greek determiners are also specified for [gender] whereas English ones are not,
- finally, features vary in their strength; for example, T in English is weak while in French it is strong resulting in a number of differences in the word order of the two languages (White, 2003).

To the above, we may add feature configurations as possible loci of parametric variation; that is, a functional category may exist in two languages but comprise of different bundles of features (e.g. Lardiere, 2009). Lardiere (2009), for example, argues that the problem for the L2 learners is the (re)assembly of features; in case these are differently assembled in the source and target languages may cause problems for L2 learners. An example of feature reassembly

---

[2] Researchers referred to (un)interpretable features in early (checking-oriented) minimalism, while more recently (in agree-based minimalism), many refer to (un)valued features. These are basically terminological differences rather than substantive (Adger & Svenonius, 2011).

[3] The EPP stands for the Extended Projection Principle (Chomsky 1982:10).

[4] Note that there is no consensus about this; whether all functional categories or only a subset of them exist in all of the world's languages.

is the aspectual distinctions encoded in past tenses in Spanish and English; both English and Spanish make use of the same features, but these are assembled differently in the two languages; that is, it is the past simple encoding habituality in English while in Spanish is the Imperfect tense and not the Preterit.

In summary, features are the building blocks that syntax manipulates to generate a syntactic derivation visualised as a tree diagram showing the hierarchical nature of clause structure to which we now turn.

2.0.3 The clause structure

In much minimalist literature the clausal structure has been described as being expressed by a C-T-V sequence of projections: [$_{CP}$ ... C [$_{TP}$ ... T [$_{VP}$ ... V]]]. VP being the lowest, embedded within TP which is higher which is in turn embedded in CP which is the highest zone (Rizzi & Cinque, 2016). According to Chomsky, these three are the core domains in a clause[5]. CP is supposed to be related to discourse, TP to inflection, and VP to thematic assignment respectively.

Figure 2.2 illustrates the core spine of a clause.

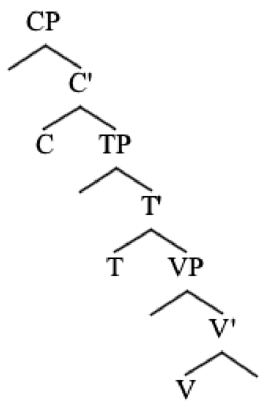


Figure 2.2: The core spine of a clause

In this work, the focus is on the Tense Phrase (TP) previously known as Inflection Phrase (IP). Before going into detail with respect to TP, let us briefly consider how the generative framework works for SLA research.

[5] vP may also be added in case of transitive or ditransitive constructions.

2.1 The generative approach in SLA

A central question for generative SLA is to what extent the same acquisition mechanisms that enable L1 are also active in L2, namely UG and parameter setting. Currently there is consensus[6] that UG is accessible but there is debate regarding the accessibility of formal features, i.e. whether UG access is full or partial (e.g. Hawkins, 2019).

Also, the learning task as far as grammar is concerned has been reworded as the acquisition/activation of the features of that language as we shall see below. More concretely, features' role to operations such as merge or agree is very important and hence despite the fact that the syntactic operations are universal, the features in an L2 are not necessarily all activated. Hence, what concerns L2 acquisition research nowadays is whether learners acquire these features as well as their interactions with the syntactic operations (Liceras, 2010).

A central open question is if L2 speakers can acquire features in the L2 especially if these do not exist in their L1. We will go over this debate in detail in the next chapter and see how different approaches account for it.

Let us now turn to the object of investigation of this study, the verbal morphology corresponding to tense, agreement and aspect features on T (and potentially a separate Aspect Functional Category) providing a generative analysis of these features.

2.2 The functional category of finiteness

We standardly take '*to sleep'* in example 2 below as non-finite and '*is'* as finite.

(2) *To sleep* enough *is* good for your health.

According to Adger (2007:3), finiteness is 'a functional category which expresses whether certain aspects of the semantics of a clause are temporally anaphoric or not, and it does so by means of a feature, which we might call [±finite].' This functional category projects independently in a particular position, i.e. the tense phrase (TP), one of the core spinal projections of a clause. T also hosts tense and agreement features but also other properties such as mood.

In English, finiteness has morphological effects, e.g. tense and agreement inflections on the verb, and syntactic effects such as the assignment of nominative case to the DP of Spec, T. Also, in languages like German, finite verbs occupy different structurally positions from non-finite verbs (e.g. Klein, 2006). Finiteness also has semantic and pragmatic effects such as

[6] Note that there are other approaches such as the usage-based or emergentist models that do not assume the existence of UG at all.

temporal deixis through the use of tense marking.

In this work, I focus on the morphological effects of finiteness on the verb forms and to the corresponding tense and agreement features on T.

2.2.1 Tense and agreement

Comrie (1985:1) defines tense as 'the grammaticalization of location in time'. In other words, tense is the linguistic expression of the location of events on the timeline and constitutes a grammatical category (Comrie, 1985) as it is encoded in a morphosyntactic way. In addition, it is a deictic category as it locates the event or situation time in relation to the speech time.

For example, consider the following sentences in English:

(3) She *is playing* the piano.

(4) She *was playing* the piano.

The difference between these two sentences is a difference in tense, the former referring to present time while the latter refers to past time with respect to the time of reference which in this case is the speech time. (Of course, (3) might be used for future reference, but we leave such use aside.)

Under current assumptions within generative syntax, tense, mood, and aspect are all contained within TP (e.g. Adger, 2003; Gelderen, 2013). The T head then hosts features such as [±finite], [±past], [±person], [±number], [±perfective], etc. The tense feature is interpretable as it contributes to the temporal interpretation of a sentence. Agreement establishes a syntactic relationship between the interpretable features e.g. the [person] and [number] features of the subject with the corresponding uninterpretable features on the finite verb of a clause.

2.2.2 Aspect: Grammatical versus Lexical

Aspect is a category intrinsically related to tense. Whereas tense locates a situation in time, aspect highlights the 'internal temporal contour of a situation' (Comrie, 1985:6). It shows how the speaker views an event or a situation: in its entirety or as being in progress.

Grammatical aspect is marked morphosyntactically through e.g. auxiliaries or verbal inflections and may be perfective or imperfective. The former is 'bounded' in the sense that it denotes the beginning and end of an event. Imperfective aspect is 'unbounded' as it does not denote any specific endpoint, instead denoting an event as ongoing. Thus, grammatical aspect describes whether a situation is viewed as complete or ongoing (Salaberry & Shirai, 2002). For example, in sentences 5 and 6 below the difference between the sentences lies in aspect, the former being perfective and the latter imperfective, and the different choice of aspect depends on the way the speaker views the situation.

(5) She *played* basketball.

(6) She *was playing* basketball.

Lexical aspect is aspect that is related to the inherent meaning of a lexical item i.e. a verb and its interaction with internal and external arguments and adjuncts (Andersen & Shirai, 1996; Salaberry & Shirai, 2002; Smith, 1983; Andersen & Shirai, 1996). For example, consider (7) below:

(7) She *is eating*.

The verb '*to eat'* is an activity verb in that eating is unbounded and durative. In example (7) it is combined with the progressive morphologically marked through the auxiliary *be* and the use of the suffix *'-ing'* on the verb stem. In combination with an activity, the reading of the progressive is that the eating is ongoing at a given reference point.

In a sentence such as 'She *ate'* which is an inherently activity verb, this becomes bound thanks to grammatical aspect. There is then interaction between lexical and grammatical aspect affecting the reading of the sentence.

Further examples clarifying the meaning of lexical aspect are given in (8) and (9): '*like'* in example (8) is inherently stative while '*run'* in example (9) is inherently atelic (without inherent temporal endpoint). However, while '*run'* can become bound as in 'she *ran* with herself' or progressive as in 'She *is running*', verbs like '*like*' which are statives are incompatible with progressive as in '*she *is liking* coffee'. Lexical aspect then is incompatible with some grammatical aspect versions and grammatical aspect can change an inherently atelic to a telic interpretation.

(8) She *likes* coffee.

(9) She *runs* with her friend.

Vendler (1967) provides a classification of verbal predicates into four aspectual classes: state, activity, accomplishment, and achievement (presented in Table 2.1) (e.g. Shirai & Andersen, 1995; Haznedar, 2007). Although there are researchers who have proposed more lexical aspectual classes (e.g. Smith, 1991) or fewer (e.g. Dowty, 1986), Vendler's classification of verbal predicates is still the most commonly used in the research community (Ayoun & Salaberry, 2008).

Each aspectual class is characterised by a different set of semantic features: [punctual], [telic] and [dynamic]. Punctual events lack duration, telic events include an inherent endpoint and dynamic events involve (internal) change (Shirai & Andersen, 1995). Table 2.2 shows the features that define each lexical aspectual class. The dynamic features distinguish states from all other classes while punctual distinguishes achievements from all other classes. Telicity characterises accomplishments and achievements contrasting with states and activities.

Table 2.1: The lexical aspectual classes and some examples

Lexical aspectual classes	**Example verbs**
state	want, love, know
activity	play, run, dance
accomplishment	run a mile, make a chair, build a house
achievement	recognize, die, find

Table 2.2: Vendler's (1967) classification of verbs into lexical aspectual classes

Lexical aspectual classes	**Semantic features**		
	Punctual	Telic	Dynamic
state	-	-	-
activity	-	-	+
accomplishment	-	+	+
achievement	+	+	+

2.2.3 Summary

To sum up, both tense and aspect are properties encoding temporality but they serve different functions. Figure 2.3 (taken from Ayoun, 2013:2) summarises how linguistic time is expressed in terms of grammar as well as lexically across languages.

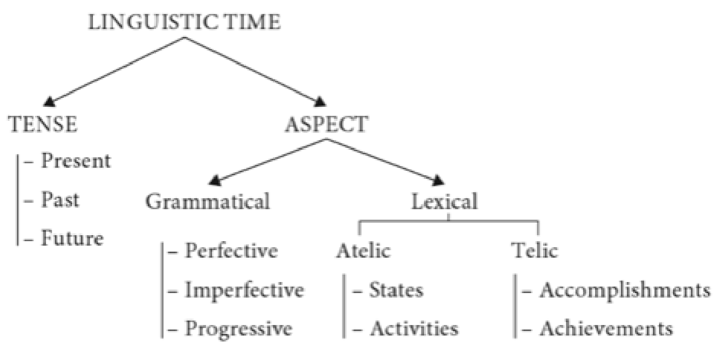


Figure 2.3: Linguistic time expressed through tense, grammatical aspect and lexical aspect

2.3 Finiteness in English, Chinese and Russian

In this section, I present the functional category of finiteness in English, Chinese, and Russian and discuss how tense and aspect are realized in the three languages respectively.

2.3.1 Finiteness in English

English finite verbs are marked for tense and/or agreement either through unbound forms such as the copula and auxiliary *be* as well as through bound morphemes such as the third singular present -*s* (3SG) and the past tense suffix -*ed*. As Table 2.3 shows, the 3SG -*s* marks the verb for agreement with a third person singular subject (e.g. she, Mary, the child) and the morpheme -*ed* marks the verb for past tense. Copula and auxiliary *be* are suppletive forms also marked for tense and agreement with the subject (e.g. he is vs he was, or I am vs they are). The suffix -*ing* marks the verb for progressive aspect while the auxiliary *be* preceding it is inflected for tense and agreement.

Table 2.3: The morphological features under investigation, their functions, and relevant examples

Morphosyntactic features	**Use:**	**Example**
-s	agreement between the T head and the DP in the Spec of TP in present tense	*She visits her family every Sunday.*
-ed	past tense	*He cooked for his friends.*

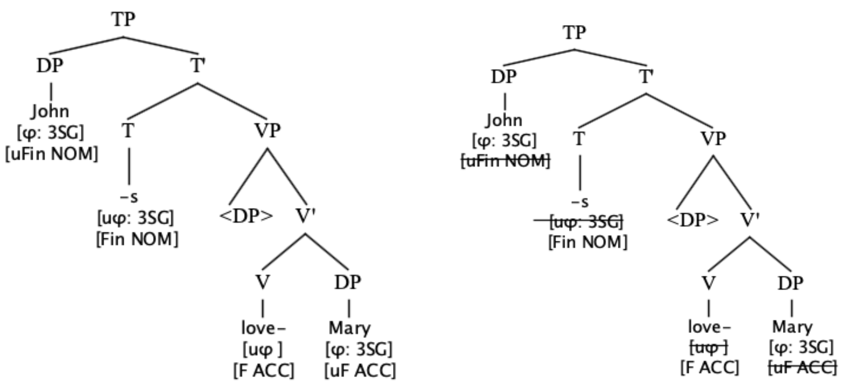
Tense and aspect are fused in English morphology. For instance, the present agreement suffix -*s* also marks imperfect (and habitual) aspect. Similarly, the past tense suffix -*ed* marks perfect aspect. Finally, the auxiliary *be* marks progressive aspect in addition to person, number, and tense and the verb inflected with -*ing*.

Table 2.4 presents both tenses and their aspectual distinctions in English along with examples.

Table 2.4: The verbal morphology system in English

English				
Time	**Present**		**Past**	
Tense	Present Continuous	Present Simple	Past Continuous	Simple Past
Aspect	Progressive	Habitual	Progressive	Perfective/ Habitual
Formation	be + V-ing	V+ -s	be + V-ing	V + (-ed)
Example	She is working.	She works.	She was working.	She worked.

Let us now see how the minimalist approach views the process of derivation of a simple transitive sentence and how it accounts for agreement and tense marking. For example, consider the sentence ‘John loves Mary’, a classic example in syntax (Koeneman & Zeijlstra, 2017). I skip here unnecessary details such as the thematic assignment within the VP and consider directly the functional layer TP. The third person singular agreement *-s* is in the T node, the position that hosts tense and agreement features. The DP-subject (‘John) encodes a bundle of interpretable features [φ: person: third, number: singular] while it also carries an uninterpretable [uFIN] feature. The T head carries uninterpretable [uφ: person: third, number: singular] and an interpretable [FIN] feature. The person and number features of D value the corresponding uninterpretable features on T and the Fin Nom feature of T values the corresponding feature of the DP which moves to Spec, TP. Figures 2.4a and 2.4b below show schematically subject-verb agreement.



Figures 2.4a-2.4b: The derivation of a simple declarative sentence in English and the workings of agreement

According to current assumptions, all thematic verbs need to raise from their original position within VP to the functional layer TP to check features of tense and agreement (Chomsky, 1995). In English, however, this raising is covert (in simple present and past tense which do not involve auxiliaries), so we do not see the result of this morphology. By contrast, auxiliary *be* forms raise overtly to T to check tense and agreement features.

The morphological merger (or M-merger) takes care of the Spell-out in morphology. The *-s* and *-ed* suffixes originate in T and through 'affix lowering' they are attached on the verb. Figure 2.5 below illustrates affix lowering.

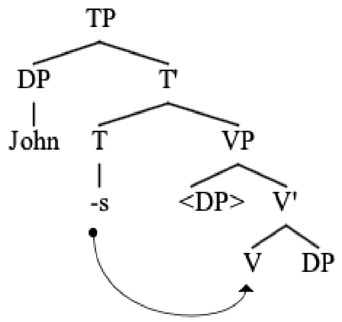


Figure 2.5: Affix lowering

2.3.2 Tense and finiteness in Chinese

Chinese encompasses languages that belong to the Sinitic language-family, a branch of the Sino-Tibetan languages. When I refer to Chinese[7], I mainly refer to Mandarin and less to Cantonese as most studies discuss the Mandarin dialect and fewer the Cantonese. Though different in many ways, all these languages are isolating, that is, they have almost no morphology.

Chinese does not overtly mark tense or agreement on the verb. It marks only aspect through particles (Li & Thompson, 1981). Thus, while Chinese lacks morphological tense and agreement, aspect in this language is more frequent yet not obligatory in discourse.

Chinese has (at least) four aspectual particles, i.e. *zai, le, guo,* and *zhe* each specified with a different aspectual value. The particle *zai* precedes the verb, while the other three particles, *le, guo,* and *zhe* follow the verb (Klein et al., 2001). The particle *zai* marks events as progressive and is usually used with dynamic verbs. The particle *zhe* marks events as durative and can be used with all verb types. The particle *guo* is an 'experiential marker' denoting an event or a

[7] In both this study and in this literature review.

situation that has been experienced at some point usually in the past but which does not hold in the utterance time. Finally, *le* is generally thought to mark completion or termination depending also on the verb is attached to. For an extensive discussion see e.g. Li & Thompson (1981) and Klein et al. (2001).

How is time reference indicated in Chinese? It cannot be denied that there are some correlations between the aspect markers and time reference. Thus, events marked for completion often will have been completed in the past (as in example 10). As a result, there are clear correlations between the perfective marker *le* and the past time reference. However, in Chinese it is also possible to use the perfective marker to indicate that an event will be completed in the future.

(10) Ta shua le liba
PR3sg. paint LE fence
He painted the fence.

As a result, Chinese uses other mainly non-morphological means to indicate time reference. These devices include temporal adverbs. Discourse factors, context and world knowledge are also important for situating the event in time, especially because none of the markers (aspectual markers or temporal adverbials) are obligatory in the language.

Smith (1991) argues that in sentences where no aspect marker is used the interpretation depends on the context and world knowledge. Smith calls the aspect in this kind of sentences 'neutral' aspect while Klein et al. (2001) call this 'zero marking'. The 'zero-marked' sentences may be interpreted in 3 ways: 1. If the sentence is part of a narration, then the narrative content will provide a hint for the right aspectual and tense interpretation, 2. they could be interpreted as imperative, and 3. they can be used to express habituality.

Example 11 below shows a 'zero-marking' sentence getting a habitual interpretation.

(11) Laoshi jiao xuesheng.
Teacher teach student.
'A teacher teaches students.'

Table 2.5 summarises the aspectual system in Chinese as presented above.

Table 2.5: The aspectual system in Chinese

Aspect	Progressive	Imperfective	Perfective	Perfective	Neutral
Formation	zai + V	V + zhe	V+ guo	V + le	V

Based on the fact that Chinese lacks overt tense and agreement marking on the verb but rather employs other linguistic means to mark tense such as temporal adverbs or contextual cues, it has often been considered a 'tenseless' language (Chen & Husband, 2018). This has raised a huge debate on whether the language in fact has syntactic tense, a Tense Phrase (Chen & Husband, 2018). Does the fact that Chinese lack overt marking mean that syntactic Tense is also absent?

While the question is unsettled, I assume that sentences are all projections of T following Adger (2003).

The question then that immediately arises, is if Chinese has syntactic tense which features are present on Chinese T? Is [past] specified on T?

Hawkins & Liszka (2003) propose that that there are obligatory and optional syntactic features, and that finite T is uniformly selected in all languages but the feature [past] is optional. Under this view, the parametric variation between English and Chinese amounts in the absence of [past] feature in Chinese[8]. This view is also adopted by Cabrelli Amaro et al. (2017).

2.3.3 Finiteness in Russian

Russian is classified as a synthetic language in terms of typology as it has rich morphology including verbs. Russian verbs are marked for both tense and aspect. There are three tenses: Present, Past, and Future and two aspects: perfective and imperfective. For the purposes of this study, I discuss only Present and Past tenses.

The majority of verbs in Russian have two forms: a perfective and an imperfective. Imperfective forms are generally 'morphologically simple, or underived verb stems' (Mezhevich, 2008: 371). The perfective forms are derived morphologically through the following processes: 1) a perfective prefix is added to the imperfective form (e.g. *pisat* 'write-IMP', ***na****pisat* 'write-PER), 2) through addition of a suffix (e.g. *kolot* 'to stub-IMP', *kol-nu-t* 'to stub-PER'), 3) through suppletion (e.g. *brat* 'to take-IMP', *vzjat* 'to take-PER') or 4) through vowel alternation (*brosat* 'to throw-IMP', *brosit* 'to throw-PER') (Mezhevich, 2008: 372). Prefixation is the most common, suffixation is used less frequently while suppletion and vowel alternation are rare.

---

[8] Mai & Yuan (2016) suggest that a [past] feature is present on Chinese T in *shi...de* constructions. As this appears a restricted domain that we are not focusing on here, I adopt the view that there is no [past] feature on T.

Further, verbs in Russian have different morphologically forms for past and non-past stems; past stems end in a vowel and non-past stems end in a consonant or glide [j] (Mezhevich, 2008).

Let us now consider tense formation in Russian. Table 2.6 illustrates the verbal morphology in Russian. The examples were taken from Mezhevich (2008).

Table 2.6: Verbal morphology in Russian

Russian			
Time	**Present**	**Past**	
Tense	Present	Past	
Aspect	Imperfective	Imperfective	Perfective
Formation	Imperfective V + suffix [3SG]	Imperfective V + suffix [Past], [SG], [Gender]	perfective V + suffix [Past], [SG], [Gender]
Examples	*Vasja čitaj***et** *knigu.* Vasja is reading/reads a/the book.	*Vasja čital knigu.* Vasja was reading/read a/the book.	*Vasja pročital knigu.* Vasja has read/read a/the book.

There is only one present tense in Russian allowing only imperfective aspect. To form it, the imperfective present form of a verb is inflected with a suffix encoding features of [person] and [number]. For instance, in example (12) below, we see the verb inflected with -*it* which denotes the present tense and encodes agreement with the subject (3SG). The post-verbal marker then expresses both tense and agreement.

(12) Stomatolog lech**it** zuby.
Dentist cure[Present 3SG] teeth.
'A dentist cures teeth.'

The reading of the present tense form can be either progressive or habitual/iterative or as stating a fact depending on the context (Bailyn, 2012).

Turning to the past tense, both the imperfective and perfective aspects are allowed with the use of the respective imperfective or perfective verb forms. To form the past tense a suffix -*l* is attached post-verbally to a past stem verb as in the example (13):

(13) On pokrasi**l** zabor.
He PERF-paint-PAST-MASC-SING fence.
'He painted a/the fence.'

The verb in (13) is marked for both [past] and [gender: masculine]. Another suffix may be added to mark the features [gender: feminine] or [number: plural]. The absence of any other morpheme means that the form is specified as masculine in gender and as singular in number. If the subject was feminine and singular, the form would need to be further marked with a suffix -*a* to mark those features. Plural is not marked for gender. Thus, verbs are marked both for past tense and agreement (gender and number agreement) with the subject.

The verb in (13) is also marked with the prefix *po*- which denotes the perfective aspect. Without the prefix *po*-, the meaning of the verb changes to denote ongoingness as can be seen in example (14):

(14) On krasil zabor.
He paint-PAST-MASC-SING fence.
'He was painting a/the fence.'

Thus, the imperfective past stem is used to denote imperfective aspect in past, and the perfective past stem is used to express perfective aspect. What is important here is that aspect and tense are encoded in different morphemes in Russian past tense.

Note that translations provided for the past tense may vary and depend on the context and the speaker's interpretation.

A few notes on the morphological realisation of tense in Russian are important for interpreting children's data; 1) Russian can omit the copula *be* in the present tense, however, in past (and future) tense the copula *be* is obligatorily used, 2) Russian lacks an auxiliary *be* (apart from the Future tense), and 2) 'Whereas English uses mainly auxiliary verbs to express both tense and aspect, Russian uses affixes for aspect and inflection for tense' (Cubberley, 2002;151).

In terms of syntax, verbs in Russian do not raise to T showing the same preferred syntactic structure as verbs in English. As a result, the structure of the English and Russian declarative clauses is the same. Figure 2.6 illustrates the phrase marker (i.e. tree diagram) of a declarative clause in Russian.

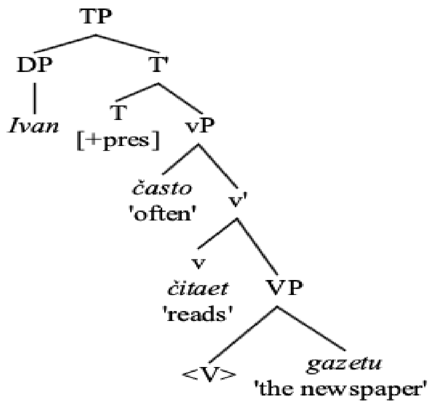


Figure 2.6: The structure of a simple declarative clause in Russian. Taken and adapted from Bailyn (2012:74)

Tense is marked on T by a tense feature such as [+pres], [+past], etc. which enters into an agree relationship with the main verb (Bailyn, 2012). Also, both present and past tense verbs are marked for agreement; present tense verbs show verbal agreement for [person] and [number] with the DP in Spec, T, while past tense shows agreement for [number] and in the singular also for [gender] (i.e. masculine, feminine, neuter) with the DP in Spec, T. The syntactic operations are the same as those described for English as they are considered to be universal. However, the feature specification in Russian differs from English in two respects: Russian encodes both gender and number agreement in past tense and aspectual features are generally separate from tense.

2.3.4 Differences and similarities between the language pairs and the learning task

The above overview of tense/aspect systems in the three languages shows cross-linguistic differences; English fuses tense and aspect in one morpheme in the past, while it marks person agreement only on 3rd person singular in present; Chinese does not encode tense morphologically but marks only aspect through (mainly) unbound particles; finally, Russian marks verbs for both tense and aspect but these are expressed through different affixes.

Formally, English and Russian realise both tense and agreement features while Chinese does not (cf. Table 2.7). These differences make the learning task different for the Chinese and Russian learners.

Chinese learners need to acquire new features not existing in their L1 and hence not activated; Russian learners can rely on the tense and aspect features activated in their L1 to acquire English verbal morphemes. It is crucial to note though that there is not one-to-one correspondence between the Russian and English forms. For example, in English there are two present tenses expressing the progressive and the habitual aspect respectively while there is only one form in Russian expressing both. In other words, from one form encoding both aspects (both progressive and habitual) in their language Russian learners will need to acquire two different tenses encoding the aspectual distinctions separately. Similarly for past, English *-ed* fuses tense and aspect in one morpheme, while in Russian there are separate morphemes for tense and aspect.

Table 2.7: Realization of tense and agreement features in the three languages

Functional category	Phenomena	Languages		
		ENGLISH	**CHINESE**	**RUSSIAN**
TP	3SG –*s*	[T], [3SG]	[T]	[T], [3SG]
Agreement feature		✓	☓	✓
TP	PAST –*ed*	[T], [Past]	[T]	[T], [Past] [gender] [number]
Past feature		✓	☓	✓

CHAPTER 3

SLA THEORIES

3.0 Child L2 acquisition: Definition

How can we define cL2? Schwartz (2004) defined L2 child as one being first exposed to a second language between the ages of 4 and 7 approximately. Her reasoning is that by age 4 the L1 grammar is mostly at place, thus making cL2 a case of sequential bilingualism, distinct from simultaneous bilingualism. The upper boundary to age 7 was based on empirical studies such as Johnson & Newport (1989) and De Keyser (2000) who found that after age 7 there is no guarantee that ultimate attainment will be native-like. These boundaries were also adopted by Unsworth (2005) as well as Meisel (2009) who provided arguments from neurological studies. However, these boundaries are not universally accepted. Genesee et al. (2004) set the lower age boundary at age 3 arguing that by this age L1 grammar is almost established. The upper boundary has been set at 9 (Penfield & Roberts, 1959), at 12 (Lenneberg, 1967), and at 15 (Long, 1990). Despite the lack of consensus within generative research the onset of cL2 is usually placed between 4 and 8, as pointed out by Haznedar & Gavruseva (2008) and Chondrogianni (2018). Following this research, I take as child L2 a child whose exposure to the L2 starts between the ages of 4 and 7 (e.g. Schwartz, 2004; Meisel, 2009).

In the remainder of this chapter, I consider the question of age in relation to the Critical Period Hypothesis and maturational effects in Section 3.2 and then turn to the question of the L1 influence over the L2 development in Section 3.3. I summarise the key points of the chapter in Section 3.4. Before so, let us briefly consider in the next Section the factors that play a role in L2 acquisition.

3.1 Factors that play a role in L2 acquisition

In this section, I consider the internal psychological and cognitive characteristics of child learners as well as external factors such as the socio-economic background of their family. I consider the role of these factors both in naturalistic and instructed settings focusing on the acquisition of grammar.

Child-internal factors

Age

Age is significant for L2 acquisition and one of the most researched factors as is argued to affect the outcome of the learning. Empirical research has given mixed results with respect to its impact (cf. Chapter 4) on the acquisition of grammar by cL2 learners with the debate within the generative framework to be revolved around the question of whether cL2 learners have the same underlying representations as cL1 or aL2 learners. It is still an unsettled issue and its impact is not yet well understood.

Aptitude

Language aptitude is a cover term for a set of cognitive abilities that predict how well an individual will learn a language compared to others in a certain timeframe and under particular conditions (Carroll & Sapon, 2002). These cognitive abilities as measured by the Modern Language Aptitude Test (MLAT), the currently most widely used aptitude test, include phonetic coding, the language analytic ability, and memory (Li, 2015). Language aptitude is a construct related to but not synonymous to the individual's general intelligence. Focusing on child SLA, Paradis (2011) found that short-term verbal memory and analytic reasoning which are cognitive abilities pertaining to language aptitude were significant predictors of accuracy in verb morphology by child L2 learners in a naturalistic setting. Sun et al. (2016), in line with Paradis (2011) showed that very young EFL children with better analytical reasoning ability may be better at receptive grammar learning. In a similar vein, Roehr-Brackin & Tellier (2019) examining English children (aged 8-9) learners of French in an instructed setting found that aptitude significantly predicted L2 achievement on grammar with analytic-language ability being the strongest in terms of predictive power. It was further shown that aptitude is dynamic in children under 12 years of age which develops along with maturity.

The typology of the L1

As L2 children have already established an L1 when they start acquiring their L2, we need to consider the influence the L1 can exert over the L2. Typological distance or proximity between an L1 and L2 can affect the influence of L1 on L2. L1 effects are well-attested in various domains of cL2. For example, Haznedar (1997), Mobaraki et al. (2008), etc. found word order transfer from the children's L1 to their L2 especially at the initial stages of acquisition. Paradis (2011) testing children (with mean exposure to English: 20 months) of various L1s on the acquisition of tense morphology in English reports differences in performance of children whose L1 marks tense morpho-phonologically and those whose L1 did not. In contrast, there are also studies such as Paradis et al. (2008) that did not find any systematic L1 effects on the acquisition of tense morphology in English by children of various L1 backgrounds during the early stages of development (with mean exposure to English: 9.5 months). These studies show (some) L1 effects in the acquisition of L2 morphosyntax, however, there is still need for further research.

Child-external factors

Quantity of input

Quantity of input is usually operationalised in terms of length of exposure (LoE), (Unsworth et al., 2011). Unsworth (2016) found that amount of exposure variable was a significant predictor for cL2 performance in the acquisition of L2 English morphosyntax (3SG and 3PL) by Dutch learners. Similarly, Paradis (2011) also found length of exposure to be a significant predictor of performance on verb morphology in English by children of various L1 backgrounds. Turning to instructed settings, Sun et al. (2016) found that amount of input at school was a significant predictor for the receptive grammar of very young Chinese EFL learners.

Quality of input

Quality of input describes the richness and nativeness of input an L2 learner is exposed to. Richness refers to the variety of input mediums (books, media, television, etc.). Jia & Fuse (2007) found that an English composite score for language environment (i.e. number of hours of TV, number of books, the number of friends, the percentage of time speaking each language) accounted for the Chinese children's performance in L2 English verbal morphology. Paradis (2011) also reported richness of the English environment (outside school) as a significant predictor of children’s performance.

Socio-economic status (SES)

The socio-economic status of a family usually measured through parental educational levels, or economic indexes such as occupation or salary range, has also been shown to have an impact on L2 language development. For example, Armon-Lotem et al. (2011) have found that maternal education is a predictor of performance in various proficiency/language ability tasks.

There are more variables to affect L2 acquisition such as the learner’s motivation, the cultural capital, the status of the L2 language which perhaps impacts on motivation, the L1 literacy, the L2 proficiency of the parents (e.g. Chondrogianni & Marinis, 2011), the personality characteristics of the learner (e.g. anxiety, extroversion/introversion, risk-taking, self-esteem), the gender of the learner, etc. It is out of the scope of the present overview to provide an exhaustive list of the factors influencing L2 acquisition. I have here introduced the variables that I will consider in my study either as main variables, age and input or background information, e.g. parental socioeconomic background.

3.2 The age factor in cSLA

The age of onset factor has been approached from various perspectives. A central empirical question is if there is an age boundary after which a learner cannot reach native-like proficiency anymore. If such a boundary exists, the question is whether it involves a cut-off point leading to a different final outcome from L1 acquisition or whether there is a gradual fading of attainment. The theoretical question is what underpins any age boundary, whether age of onset effects are due to maturation of a domain specific linguistic faculty or the result of other individual and contextual factors such as motivation, identity, quantity and quality of input (De Keyser, 2013).

This thesis aims to address the empirical question of the potential impact of age of onset focusing on populations within childhood. In what follows, I review the relevant literature on age effects in L2 acquisition.

3.2.1 Age effects and the Critical Period Hypothesis

This idea that there is a critical period for language acquisition, that is, someone who starts learning a language after puberty is unlikely to reach native-like levels is old and goes back to Lenneberg (1967). Lenneberg argues that the end of the critical period is before puberty at a time when the maturational processes in the brain have been completed having reached the adult values and resulting in 'loss of adaptability and inability for reorganization in the brain' (Lenneberg, 1967:179). Completed neurological processes in the brain mark the end of the critical period with the boundary set at around 10-12 years of age.

Lenneberg talked about the Critical Period Hypothesis in relation to first language acquisition. With respect to second language acquisition, he only suggested that learning a second language automatically becomes increasingly difficult after puberty or even impossible especially if this happens by mere exposure to input because of the completion of the maturational period.

The Critical Period Hypothesis for second language acquisition was empirically tested in an influential study by Johnson and Newport (1989). They tested 46 adult Korean and Chinese immigrants in the U.S. who had different ages of onset in learning English as determined by their age of arrival. The participants were tested on a number of features and structures in a Grammaticality Judgment Task (GTJ) and results showed a strong negative relationship between age of arrival and task performance, i.e. the older the age of arrival of the participant and hence their age of onset, the worse their performance. In particular, they found that L2 learners with age of onset between 3-7 would perform as native speakers, L2 learners with age of onset between 7 and around puberty would show a decline in performance while for those L2 learners with age of onset after 17, age of onset was not a factor predicting their performance anymore. In order to further test possible confounding with age variables, they also measured experiential (such as length of exposure and amount of initial exposure) and attitudinal (such as motivation and identification) variables. The former did not seem to significantly correlate with L2 learners' performance. The latter did correlate with L2 learners' performance but were not found to account for the observed differences as well as age of onset did. The authors then argued that the critical period hypothesis can be extended to include second language learning and gave support to the maturational account as its explanation which Lenneberg (1967) had proposed and according to which a language needs to be acquired within the maturational period in order for learners to fully acquire it. In contrast to Lenneberg though, they found that the age boundary is earlier than the one he proposed lowering it to the age of 6-7. Furthermore, in contrast with what Lenneberg's hypothesis suggested, they claimed that this age boundary is followed by a gradual decline in ultimate attainment for those with AoO between 7 and 17 rather than an abrupt change in performance.

Hyltenstam & Abrahamsson (2003) though reviewing previous studies showed that most of them converge on the fact that even a short delay in exposure to the L2 can result in non-native-like attainment. They thus suggested that instead of a cut-off point at a certain age full native-like ultimate attainment is fading as an effect of an increasing age of onset. Maturational effects then are manifested in the linear decline of performance as age of onset increases. Maturation interacts with other variables that affect acquisition such as social/psychological factors. Reversely, these factors become more important with age.

More recently, Meisel (2009) based on neurological studies -which use electroencephalography (EEG) and haemodynamic methods and show that the activation patterns as well as the spatial organization of the brain are different between first language learners (monolingual or bilingual) and second language learners especially when learners are exposed to syntactically ill-formed sentences- proposed an even lower age boundary for the acquisition of grammar by L2 children. These changes are mainly observed around the ages of 4 and then 6-7. Drawing also from linguistic evidence -such as results from studies on the acquisition of morphosyntax by early L2 learners showing qualitative differences in their performance from L1 children- he proposed that the age 4 is a possible dividing line between (bilingual)[9] first language acquisition and cL2. This age boundary may be followed by a gradual decline in performance. The implication of his proposal is that a child that starts learning a second language before 4 will pattern as monolingual or simultaneous bilingual children (2L1s). In contrast, for a child that will first be exposed to the second language after 4, their acquisition of grammar is expected to be problematic and will pattern similar with L2 adults.

3.2.2 Age and Representations

Various theoretical hypotheses have been put forth to explain how language learning takes place after the critical period, if it involves a different learning route, if UG is accessible. One key empirical difference between L1 and L2 acquisition is the morphological variability found in L2 learners' performance which has ignited the debate regarding the nature of syntactic representations in L2 acquisition. The question is if the optionality in morphological marking observed even in advanced L2 learners is due to an impaired syntactic representation or due to problems of mapping syntax to morphology.

All possible theoretical hypotheses have been considered in relation to access to UG post-critical period, namely no access, partial access, and full access. Bley-Vroman's (1990) Fundamental Difference Hypothesis (FDH) argues that L2 learners cannot access the UG anymore and have to make use of other cognitive systems. This hypothesis does not find many proponents nowadays as much empirical evidence shows that L2 is constrained by UG (e.g. White, 2003; Meisel, 2011; Mitchell et al., 2013).

Currently the debate is between the Full Access and Partial Access. The Full access to UG hypothesis suggests that the same cognitive mechanisms underlie both first and second language acquisition (e.g. Schwartz & Sprouse, 1996; Haznedar and Schwartz, 1997; Prévost & White, 2000).

---

[9] It is widely accepted by researchers in the field of bilingualism that acquisition of two languages from birth can be considered as an instance of bilingual first language development (Meisel 2009).

Partial access hypotheses suggest that L2 grammars are UG-constrained but learners will not be able to reset all parameters since formal features not activated through the L1 are not available, leading therefore to representational deficits (e.g. Hawkins & Casillas, 2008; Tsimpli & Dimitrakopoulou, 2007).

In the following sections I review these hypotheses with reference to the acquisition of finiteness, tense and aspect morphology.

3.2.2.1 Representational Deficit Hypotheses

Representational Deficit Hypotheses assume that the differences attested between L1 and L2 learners are due to maturational/critical period effects leading to different language representations. These hypotheses assume partial availability to UG which leads to a syntactic deficit. The two most influential hypotheses are the Interpretability Hypothesis (Tsimpli & Mastropavlou, 2007; Tsimpli & Dimitrakopoulou, 2007) and the Failed Functional Features Hypothesis recently revised as Contextual Complexity Hypothesis (Hawkins & Casillas, 2008). Below I review these hypotheses in chronological order.

3.2.2.1.1 Failed Functional Features Hypothesis

Hawkins & Chan (1997) proposed the Failed Functional Features Hypothesis according to which there is partial access to UG. L2 learners in post-critical-period will not be able to acquire features related to functional categories if these do not exist in their L1 and hence not activated early in life. The window of accessibility to those features is restricted to a certain timeframe, sometime during childhood. However, their grammars will continue to be constrained by principles of UG which are fully available. Their grammars then will be ‘possible grammars’ but diverging from both the target L2 and their L1s.

Hawkins & Liszka (2003) tested this claim by examining oral production of the past tense morpheme by Chinese, Japanese, and German learners of English. Chinese performed lower than the other two groups in supplying the tense morpheme which was interpreted as due to the inaccessibility of this feature for Chinese learners whose L1 lacks it. According to this account then all functional features which are not instantiated in the L1 will be inaccessible in post-critical L2 acquisition.

3.2.2.1.2 Interpretability Hypothesis

Tsimpli & Mastropavlou (2007) and Tsimpli & Dimitrakopoulou (2007) proposed the Interpretability Hypothesis; specifically, they propose that while interpretable features are acquirable in L2 acquisition regardless of their existence in the learners' L1, uninterpretable features not activated early in life in the L1 will be inaccessible to L2 learners. The logic is that interpretable features are LF-interpretable having a dual status: a linguistic and a conceptual one. They can thus be accessed either from cognition to language or from language to cognition. However, uninterpretable features which are purely syntactic and have no semantic import will remain unavailable to L2 learners.

Tsimpli & Dimitrakopoulou (2007) investigate the acquisition of wh-questions by intermediate and advanced Greek learners of English and show that uninterpretable features such as (subject, object) agreement cause problems for L2 learners. Learners compensate for these difficulties with the use of interpretable features like [animacy].

3.2.2.1.3 Contextual Complexity Hypothesis

While the Interpretability Hypothesis mainly focuses on ultimate attainment of learners and in grammatical phenomena in general, the Contextual Complexity Hypothesis (Hawkins & Casillas, 2008) focuses more on early L2 speech in an attempt to account for the performance of early L2 learners on verb morphology. Some common findings of early L2 production of English verb morphology include the more frequent suppliance of auxiliary *be* forms than affixal morphology, the omission of inflections such as the 3SG *-s* and past *-ed*, the few mismatches in case of use of an inflected verb, the overgeneration of *be* in contexts that is not required and the overgeneration of progressive forms.

Hawkins & Casillas assume that L2 learners differ minimally from native speakers as their grammars are organised in the same way due to the architecture of the language faculty: UG. However, they differ minimally in the way they store Vocabulary entries. Based on the hypothesis that L2 learners do not initially have access to uninterpretable features but do have access to interpretable ones, they propose that learners assign interpretable features to phonological strings. The distribution of the different forms in L2 speech is then due to the way phonological exponents are stored which differs from L1 acquisition. While L1 speakers have Vocabulary entries specified in terms of bundles of features at the point of lexical insertion as in the example below, L2 learners have Vocabulary entries for exponents driving dependencies which are context-sensitive:

Native speaker Vocabulary entry:

/s/ $\leftrightarrow$ [V, -past, +sing, 3p] + ___

L2 speaker Vocabulary entry:

/s/ ↔ /[V]+ ___ / [T, -past] ___ / [N, +sing, 3person] ___

In other words, L2 learners will have something for 3SG *-s* along the following lines: 'insert /s/ in the context of a verb which is in the context of a non-past T, itself in the context of a 3rd person, singular N' (Hawkins & Casillas, 2008:602).

Further, the more sister nodes (i.e. co-occurring syntactic terminal nodes) required to specify the context, the higher the difficulty in retrieving the form. To exemplify this, they take a bare verb and an inflected form. The entry of a bare verb makes no statement about sister terminal nodes; on the other hand, an inflected verb such as a verb inflected for 3SG or past tense is specified with more than one sister nodes as shown below:

/s/ ↔ /[V]+ ___ /[T, –past] ___ / [N, +sing, 3person] ___

/d/ ↔ /[V]+ ___ / [T, +past] ___

The different frequencies of verb forms produced by L2 learners then is explained by the number of sister nodes; which are more in case of 3SG *-s*.

Under this proposal, the production of *be + verb* follows as a possibility as the statement of the contexts of insertion of forms of *be* refer only to T and the properties of the subject. However, inflected with /s/ and /d/ verbs are not expected to follow *be* as these are highly specified.

Finally, to account for the difference in frequency of suppliance between copula and auxiliary *be*, the authors argue that this stems from the fact that learners come to realise that auxiliary *be* needs to accommodate an interpretable progressive feature when preceding a V-ing form. Thus, auxiliary *be* has one more sister terminal node than copula *be* which does not realise any aspectual distinction.

Over time, learners who have the relevant features in their L1 will be able to restructure their L2 representations and include uninterpretable features; learners who cannot access the features in their L1 will not be able to restructure their L2 representations showing persistent optionality in advanced levels.

3.2.2.1.4 Underspecification of AspP Hypothesis

A last approach to discuss which attempts to explain the optional infinitives in second language acquisition is the Underspecification of AspP Hypothesis proposed by Gavruseva (2002; 2003;

2004). This account links the root infinitives with the aspectual properties of the verbs. The aspectual features are viewed as key elements of finiteness in this proposal.

More concretely, Gavruseva uses two features to distinguish between verbs: telicity and punctuality. Stative verbs (e.g. know) are inherently atelic ([-telic]) and punctual eventives, that is, achievements (e.g. catch) are inherently telic ([+telic]). The telicity of non-punctual eventives (i.e. activities and accomplishments, such as 'play (a basketball game))' is dependent on the nominal arguments that may follow them; they can be [±telic). The verb's telicity semantics is claimed to determine the finiteness of the predicate.

Telicity is claimed to be a syntactic feature that needs to be checked by the verb in an AspP as shown in Figure 3.1 below (taken from Haznedar, 2007:392).

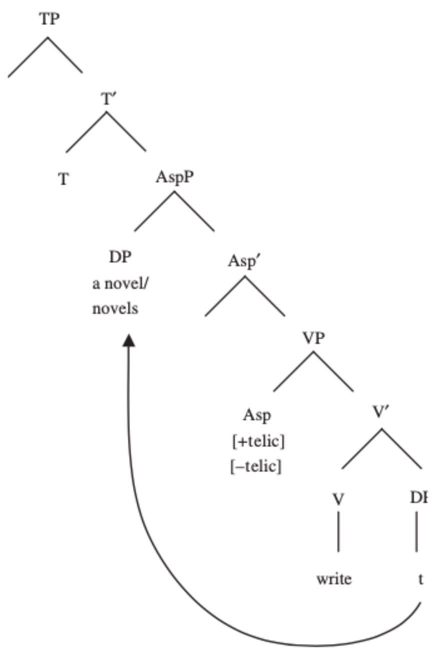


Figure 3.1: The workings of the telicity feature.

Gavruseva suggests that as states and punctual eventives are specified as inherently atelic and telic respectively (that is, telicity is intrinsic to the verb's lexical meaning), only non-punctual verbs are underspecified for telicity and must move through an AspP projection; their telicity feature will be determined by the argument (/prepositional complement/adjunct) in the specifier of the AspP.

She thus predicts that:

1. statives will be finite
2. punctual eventives will be finite

3. non-punctual eventive verbs (activities and accomplishments) will be non-finite as they remain unspecified for syntactic telicity in early child grammars. In other words, whether a feature is inherent or non-inherent (a)telic will determine whether a predicate will be finite or non-finite in children's productions.

In later work, Gavruseva (2008) also considers the acquisition of copula and auxiliary *be*. She claims that copula *be* is free of syntactic aspectual specification, whereas the auxiliary be is specified for the aspectual feature [-bound]. Under the Underspecification of AspP hypothesis, early in L2 acquisition, syntactic aspectual features are underspecified, hence the observed asymmetry between copula *be* which shows a very high production rate early on and auxiliary *be* which develops more slowly.

Finally, she considers the overgeneration of *be*. Overgeneration of *be* involves overuse of *be* in declarative contexts where affixal inflection is required. *Be* also appears as a dummy auxiliary in negative or interrogative sentences in place of e.g. *do*. Gavruseva argues that this overgeneration of *be* is due to two factors: 1) the acquisition of the copula *be* (both in morphological and syntactic terms) and 2) the misanalysis of input, that is, children misanalyse contracted 's forms as instances of *be*. As a result, children take *be* to be a generic finiteness marker using it to different rates and for different time periods. Further investigating the aspectual properties of the overgenerated *be*, she found that it mainly occurs with statives and punctual verbs, that is, achievements. Hypothesising that utterances with overgenerated *be* are finite and given that syntactic tense chains are much easier established for statives and achievements because they are inherently specified for telicity, it is entailed that VPs must be specified for aspectual features. The fact then that children overgenerate *be* in these contexts shows that the aspectual system in their grammars is emerged and influenced by lexical semantics.

3.2.2.2 Full Representation approaches

Full Access Approaches hold that L2 learners have full access to the universal inventory of features as L1 children (Schwartz & Sprouse, 1996). Variability in L2 learners' performance is not due to critical period effects. Rather, L2 learners have intact representations and the problem arises in the mappings between morphology and syntax or is due to phonological and/or processing issues. In other words, L2 learners mainly face production issues while comprehension may not be affected at all. L2 learners do not differ qualitatively from L1 learners though quantitative differences are to be expected. This hypothesis predicts omission errors as a result of failed mapping with morphology, but few if any commission errors.

Further, native-like achievement is possible under this approach. Let us see two concrete proposals within this view.

3.2.2.2.1 Missing Surface Inflection Hypothesis

According to the Missing Surface Inflection Hypothesis (MSIH) developed among others by Haznedar and Schwartz, 1997; Lardiere, 1998; 2000; Prévost and White 2000; and White, 2003, variability in performance with functional morphology does not reflect defective or impaired underlying representations but it rather shows difficulties with the overt realisation of morphology.

The argument is that a functional category requires knowledge of 3 types: morphophonological, syntactic, and semantic. They point out that syntactic evidence e.g. regarding subject realisation indicates that the functional category T has relevant finiteness features and it is only the morphological realisation of these features that is affected, indicating a dissociation between syntactic representations and morphological realisation.

Under this account, bare forms may in fact be used as finite forms. L2 learners do distinguish between finite and non-finite forms and finite forms do not substitute non-finite forms as are highly specified. Variability, under this account then, is constrained and is manifested always unidirectionally. The substitution of finite forms by non-finite forms happens due to 'mapping problems' between morphology and syntax under communication and/or processing demands. That is, although learners fully represent the functional categories and the features, they have problems accessing the surface morphological realisation of lexical items, using sometimes what Prévost & White (2000) call 'defaults' or 'elsewhere' forms.

3.2.2.2.2 Prosodic Transfer Hypothesis

The Prosodic Transfer Hypothesis developed by Goad et al. (2003), Goad & White (2004) attempts to explain the failure of L2 learners to consistently use functional morphology proposing L1 prosodic constraints; that is, when the prosodic structure of a certain phenomenon in the L2 differs from the respective structure in the L1, then learners may not use the morphophonological forms in L2 production. Under this view, perception is supposed to be intact as L1 prosodic constraints do not affect it; only production is affected. Similarly to the MSIH, L2 learners have intact representations with L1 prosodic constraints to be proposed as one of the problems L2 learners face. According to the original proposal, restrictions of the L1 prosodic constraints are always present even at advanced stages of the L2 learning. Goad & White (2006) revised this position by proposing that minimal adaptation and restructuring of the L1 licencing relations in the prosodic domain is possible, not though when the L1 lacks the structures all together. In the case of a Turkish speaker of English, for example, no restructuring will be possible in case of articles and problems will be persistent because Turkish language lacks articles.

To give another example, Chinese learners whose mother tongue does not inflect verbs for tense or agreement and no adjunction is allowed in the prosodic word - consonant clusters are generally not permitted as in English past tense forms (Cabrelli Amaro et al., 2018)- may have problems with supplying verbal inflection in English as the prosodic structure between the structures of the two languages differ. Note though that under this view, we would expect irregular verbs to be used unproblematically as the past tense morpheme is organised PWd (i.e. prosodic word)-internally – as is the case of the Mandarin prosodic shapes of functional morphology, e.g. aspectual morphemes (Goad & White, 2006).

With the same underlying assumptions as the MSIH, this is then a proposal which is restricted to explaining the bare forms used in finite contexts but it does not account for any other mismatches and errors in L2 learners' performance.

3.2.2.3 Feature Reassembly Hypothesis

Finally, another recent account for the variable use of morphology in L2 acquisition by Lardiere (2007; 2008) proposes a shift of focus from considering the presence/absence of features in two languages to the way features are assembled in different language pairs. According to this approach, the learning task for an L2 learner is the re-assembly of the feature configurations of a given category in their L2 if this differs from their L1.

To exemplify how this proposal works, Lardiere, in a case study, examined Patty, a Mandarin and Hokkien Chinese L1 speaker having been resided in the US for more than 10 years the first time of testing and 18 years in total at the second round of testing. She investigates the acquisition of number and definiteness. Plurality exists and is marked in English on count nouns, definite or indefinite. Chinese expresses plurality through the use of quantifiers preceding a noun but without overt morphological marking of plurality on the noun. Chinese though selects the [+plural] feature, but this is realised very differently from English. Specifically, the suffix -men in Chinese which is used with a specific type of nouns (i.e. denoting humans) is interpreted as plural and must be read as definite -when realised as it is optional. The learning task then for Patty or any Chinese speaker is to detach definiteness from plural marking and reassemble the features as they are in the target L2, English.

This proposal may be tested empirically if, for instance, we test Russian or Spanish learners of English whose L1s do have the relevant features and indeed in some cases are assembled differently. It is not directly visible though how it accounts for the case of, for example, Chinese learners of English whose L1 lacks tense and agreement features altogether and the learning task does not seem to be the reconfiguration of them.

3.2.2.4 A functionalist approach: The Aspect Hypothesis

From a different theoretical framework, another approach has been proposed to account for the acquisition of verb morphology; the Aspect Hypothesis (e.g. Shirai, 1991; Andersen and Shirai, 1994, 1996; Shirai and Andersen, 1995; Bardovi-Harlig and Bergström, 1996; Bardovi-Harlig, 1999). This is a proposal developed by functionalist researchers who argue that the construction of meaning guides language development. In the case of temporality, this provides the ground for investigating the associations between meaning and form; more concretely, it has been suggested that lexical aspectual properties of the verbs which are part of their core meaning guides the acquisition of tense and aspect markers.

The Aspect Hypothesis was first proposed to explain the performance of L1 learners in the verbal domain (e.g. Andersen and Shirai, 1996) and was then extended to account for performance on verb morphology by L2 learners. This proposal which states that L2 learners in early stages of learning use verbal inflection to mark aspectual properties of the verbs -not deictic tense- revolves around four hypotheses:

1. Learners first use (perfective) past marking on achievements and accomplishments, eventually extending use to activities and statives.
2. In languages that encode the perfective/imperfective distinction, imperfective past appears later than perfective past, and imperfect past marking begins with statives, extending next to activities, then to accomplishments, and finally to achievements.
3. In languages that have progressive aspect, progressive marking begins with activities, then extends to accomplishments and achievements.
4. Progressive markings are not incorrectly overextended to statives. (Bardovi-Harlig, 2000: 227)

3.3 L1 influence in cSLA

We now turn to an interrelated with age issue, the typological impact of the learners' first language on their second, that is transfer. L1 to L2 transfer can be thought of as a special case of crosslinguistic influence, the influence exerted between the languages of a bilingual speaker, e.g. including L2 to L1 (L1 attrition) or crosslinguistic influence in a simultaneous balanced bilingual. For the purposes of the current study, I adopt the term transfer as I focus on influence of L1 on L2.

Several proposals have been put forward to account for transfer effects especially considering aL2 acquisition, reviewed below.

3.3.1 Full Access Approaches and the view of transfer at the initial stages

A central question is how much transfer if any is present at initial stages of L2 acquisition. Three hypotheses have been proposed arguing for full transfer, no transfer and weak transfer. Specifically: 1) the Full Transfer/ Full Access Model (Schwartz & Sprouse, 1996), 2) the Full Access/ No Full Transfer hypothesis that could be exemplified by the Minimal Trees Hypothesis or Organic Grammar in more recent work (Vainikka & Young-Scholten, 1996a; 1996b; 2007; 2011), and 3) the Full Access/Weak Transfer hypothesis (Eubank, 1993/1994).[10]

The Full Access/Weak Transfer (exemplified by the Valueless Features hypothesis proposed by Eubank (1993/1994, 1994, 1996)) states that both lexical and functional categories are transferred from the L1 but not the feature values associated with the functional categories, that is, feature strength is not transferred and features are initially inert or valueless. Recall from chapter 2, section 2.0.2 that feature strength entails differences for word order. After this initial stage of feature inertness, feature strength may be acquired depending on the acquisition of inflectional morphology. Morphologically rich paradigms will trigger strong Infl, poor paradigms otherwise. However, as we do not look into word order differences related to feature strength, this proposal will not be considered further.

Let us see the remaining two proposals in more detail.

3.3.1.1 The Full Transfer/Full Access (FT/FA) model

The Full Transfer/Full Access (FT/FA) model proposed by Schwartz and Sprouse (1996) states that the initial state of the L2 acquisition is their L1; both lexis and grammar (i.e. functional categories, features, and feature strength). L2 learners will assume their L1 grammar until revision occurs through exposure to the target language input and subsequent setting of the TL values. This predicts that learners of different languages will show L1 effects from initial stages before restructuring occurs. L1 effects will disappear over time since learners have Full Access to features and will be able to restructure their grammars and so problems with morphosyntax

[10] There is a fourth possibility: No Transfer, which has been shown empirically incorrect (Meisel, 2011).

may not persist to advanced levels. For example, we would expect L2 learners of English from isolating language backgrounds such as Chinese where tense and agreement are not marked on the verb to show a different performance from learners whose L1s has rich inflectional verb morphology paradigms. The initial stages are expected to differ between them, while convergence is expected with increasing proficiency (e.g. Paradis et al., 2008).

3.3.1.2 Full Access/ No Full Transfer

The Organic Grammar Hypothesis proposed by Vainikka & Young-Scholten (2007, 2011) claims that at the initial state there is only lexical material organised in a linear fashion as in the learners' L1. Crucially, functional categories are not transferred; they are entirely absent at the initial stage and are built gradually over time through exposure to input. The VP will precede IP which will precede CP. The development of functional categories is totally independent of the relevant categories in L1, thus no transfer is expected at any stage of development. A prediction of this hypothesis is that there will be no discernible L1 effects; for example, Chinese and French learners' development of verb morphology will be the same at all stages of development; both types of learners are expected to attain native-like grammars under the assumption of full access to UG.

3.3.2 Partial Access Approaches and their view of transfer

Partial access to UG approaches predict a discernible L1 effect since features of inflectional morphology will not be accessible to L2 learners if absent from their L1. Thus, L1 set values determine whether L2 feature values will be acquired or not. It follows logically that inaccessible features will remain so throughout all stages of L2 development and will not be acquired. L2 learners will compensate for feature inaccessibility by resorting to either L1 feature values or options available by UG such as interpretable features.

Hawkins & Casillas (2008) also suggest that only those learners who have activated the uninterpretable features in their L1 will be able to acquire them.

3.3.3 Summary

To sum up, Full Access approaches agree that over time and under exposure to input learners will converge to target L2 values, hence any transfer effects are expected at the initial stages of L2 development. The Full Transfer/Full Access model assumes that functional morphology will be transferred from the L1, predicting that learners from inflecting backgrounds will perform better than learners from e.g. isolating backgrounds when acquiring a language like e.g. English. The Organic Grammar hypothesis argues that no transfer of functional categories occurs at any stage of L2 development.

Partial access to UG approaches expect transfer effects throughout all stages of development, but especially at later stages when learners from language backgrounds that do not have the uninterpretable features that exist in the L2 will not be able to acquire them showing persistent optionality in contrast to learners whose L1 has these features. Transfer is seen as one compensatory mechanism for the unavailability of features while use of interpretable features is another option for learners.

In the next chapter, I will present empirical findings from studies investigating the acquisition of finiteness by child L1, adult L2, child L2 learners.

CHAPTER 4

THE ACQUISITION OF INFLECTIONAL MORPHOLOGY BY L1 AND L2 LEARNERS: PREVIOUS STUDIES

4.0 Introduction

In this chapter, I review empirical evidence regarding the acquisition of finiteness (i.e. tense and agreement marking) by three types of learners: child L1, adult L2, and child L2 learners. I discuss studies on cL2 acquisition in both naturalistic and instructed settings to consider the potential impact of input/context.

The discussion begins with L1 children in Section 4.1 followed by adult L2 learners in Section 4.2 and cL2 in Section 4.3. In Section 4.4 I consider the emerging picture from all types of learners and settings before I turn to transfer effects in Section 4.5 and conclusions in Section 4.6.

4.1 The acquisition of finiteness by L1 English-speaking children

From their first two-word combinations, English L1 children often omit inflection in main thematic verbs such as 3SG *-s* and past tense *-ed* when is actually required (cf. example 1). In addition, they sometimes omit the copula and auxiliary *be* (cf. examples 2-3), as well as the auxiliary *do*.

(1) Sarah (2;3): *Marie go.*

(2) Sarah (2;7): *You (are) nice.*

(3) Eve (1;7): *Eve (is) cracking nut.*

(Brackets show missing elements. Examples are taken from Guasti (2016)).

This stage has been described as the Optional Infinitive (OI) (Wexler 1990, 1992, 1994) or Root Infinitive (RI) stage (Rizzi, 1994) and lasts from around 2 to 3;6 years of age.

This stage is further characterised by 1) regular omissions of subjects (cf. example 4) which are not licenced in English as is a non-null subject language (Hyams, 1986; Wexler, 1990; Wexler, 1998), 2) variable assignment of pronoun subjects with a default accusative case (cf. example 5); this holds especially for the cases when a non-finite verb is used (Wexler, 2002), and 3) target word order (cf. example 6) -English has a weak V- feature so it always remains in the same position regardless of whether there is a finite or non-finite form (Poeppel and Wexler,1993;Wexler, 1994).

(4) *no want stand head*

(5) *him like ice cream*

(6) *He no bite you.*

(examples from Roeper & Rohrbacker, 1994; Wexler, 1998; and, Klima & Bellugi, 1966 respectively).

While L1 children use finiteness marking variably, they nevertheless, rarely make commission errors (Poeppel & Wexler, 1993; Rice et al. 1995; Harris & Wexler, 1996) such as **They plays*, or **She are playing*. Furthermore, L1 children respect word order in case of negation; for example, they will not say **He not is playing* (Wexler, 1998).

What further characterises their acquisition of finiteness in early speech is the observation that a correlation exists between the development of morphemes, that is, finiteness morphemes are developed simultaneously following similar rates regardless whether they are suppletive or affixal. (see Brown, 1973; de Villiers & de Villiers, 1973; Rice et al., 1995; Rice & Wexler, 1996; Rice et al., 1998; Zobl & Liceras, 1994).

Finally, it is important to consider whether early child speech is also characterised by the *be*-overgeneration - a phenomenon which has been found in L2 acquisition - which is the use of auxiliary *be* in contexts that is not required (e.g. example 7 below). There are somewhat mixed results to this respect. Some researchers argue that the phenomenon of *be*-overgeneration appears to be unique to L2 learners (Paradis et al., 2004; Ionin & Wexler, 2002; Pierce et al., 2012).

However, Brown (1973) attested a few instances of *be*-overgeneration produced solely by Adam, a very young English-speaking child who uttered sentences such as '*It's went*', '*It's will go'* as well as '*It's truck'*. After careful analysis, Brown suggested that '*it's*' was an unanalysed form for Adam, so it was used instead of '*it'*. This instance then cannot be taken as a case of *be*-overgeneration.

More serious is evidence from the study of Tesan & Thornton (2004). They tested three L1 English children of approximately 1;9-3;0 years old and found that all children overgenerated *be* as in examples 7 and 8 to various degrees (ranging from 20% to 80% of all their utterances). They found in total 173 instances of this type out of a total of 1553 declarative sentences; *be* was used with the pronominals: *he, she, it,* with proper names, with various nominal expressions as well as with quantificational elements. The *be*-overgeneration did not precede use of finite forms meaning that it did not constitute an earlier developmental stage. The authors argue that it is realisation of verbal agreement.

(7) June s eat pizza.

(8) Pooh s likes pizza.

This phenomenon was also attested in case of negative sentences (N=188) as in examples 9 and 10. These constituted 10 and 11 sentences respectively out of the total 188 negative sentences.

(9) He s not eat carrots.

(10) He s don't like carrots.

Tesan & Thornton argue that as *do*-support has not emerged yet, the use of *be* realises the formal values of the relevant functional categories, that is, *be* is used instead to satisfy the *do*-support constraint.

Dye et al. (2004) claim that children do sometimes relocate -*s* bringing up examples from Santelmann et al. (2000;2002), e.g. '*Bunny's touch a carrot'* instead of '*Bugs Bunny touches a carrot'*, and data from Cornell Language Acquisition Lab, e.g. '*he's try to bite her*' and '*she's find a fork*'. Dye et al. argue that the fact that children relocate -*s* is an indication that children do have the functional category I.

There is some evidence then that L1 children do sometimes overproduce *be* forms which has been attributed to two reasons: either children are in the very early stages and have not yet analysed the input as is the case of Adam, or children relocate -*s* or use it instead of e.g. *do* to mark syntactic positions. However, to the best of my knowledge, no study in L1 acquisition has shown use of '*be + verb*' to be used with various meanings or overgeneralisation of progressive tense to various contexts, facts that characterise L2 acquisition as we will see below.

With the end of the RI or OI stage, children use only finite forms in main clauses and subjects are used in an adult-like fashion assigned nominative case.

In sum, English monolingual children go through an optional infinitive stage characterised by variable use of finite morphemes, drop of subjects, assignment of accusative case to pronoun subjects preceding non-finite verbs, and target-like word order of their language. They make very few commission errors, that is, they do not overgeneralise inflections, however, they sometimes use *be* in contexts that is not required. Tense morphemes seem to develop at similar rates and once children master their use, adult-like utterances are to be expected. Within 3;6- 4;0 years of age English L1 learners have acquired the inflectional morphology of their language.

There have been various hypotheses trying to explain the above data such as the Maturation Hypothesis (e.g. Radford, 1990) and the (Strong and Weak) Continuity Hypothesis (e.g. Poeppel & Wexler, 1993; Clahsen, Eisenbeiss & Vainikka, 1994) from a generative perspective as well as the (Primacy of the) Aspect Hypothesis (e.g. Shirai, 1991; Andersen and Shirai, 1996) from a functionalists' perspective. It is out of the scope of the present chapter and this thesis to discuss these theories in detail since this work cannot contribute to that debate. However, the empirical findings can provide some reference for comparisons with L2 children.

4.2 The acquisition of finiteness in English by L2 adults

In contrast to L1 learners who within 4 years master the inflectional morphology of their language, studies investigating adult L2 acquisition converge on that learners show variable and low accuracy on the verbal domain, that is, low suppliance of inflectional morphemes even at advanced stages and after many years of exposure. Adult L2 learners very frequently omit the 3SG -*s* as well as the past tense marker -*ed* in their speech (cf. Table 4.1).

Most research within the generative framework has focused on the acquisition of finiteness in ultimate attainment, that is after many years of exposure to the L2.

Hawkins & Liszka (2003) aiming to explore the source of difficulty of L2 learners of English to mark past tense tested Chinese, Japanese and German advanced learners of English. They considered the finite T to be uniformly selected in all languages while the feature [+/-past] to be optional, arguing that it is selected in English but not in Chinese. Crucially, the feature [+/-past] is selected by Japanese and German languages. Results of a written morphology test showed high performance for all groups. However, results on spontaneous data revealed differences between the three groups with Chinese learners of English showing significantly lower performance from Japanese and German learners, who were indistinguishable. The authors argue that the difficulty of the Chinese speakers to mark past tense is due to the fact that the syntactic (formal) feature (i.e. [+/- past]) was not activated early in life in L1 and thus, is inaccessible in L2 (due to critical period effects).

Lardiere (1998a, b) tested Patty, a Chinese and Hokkien-speaking immigrant in the US. Patty moved to the US at the age of 22 where she studied at the university at a bachelor's and master's level. The first recording session was held 10 years later, when Patty was 32, while the second and third around 8 and a half years later. Her language environment had been quite mixed in the first recording, but later on she got married to an American man and worked in a company being totally immersed in an English-speaking environment. Lardiere counted the suppliance of 3SG -*s* and Past -*ed* in obligatory contexts. Agreement marking on thematic main verbs was supplied around 4% out of all obligatory contexts while past tense marking was found to be supplied around 34% in obligatory contexts. Suppliance of nominative case as well as accusative case marking was perfectly supplied (100%) in obligatory contexts while Patty had also acquired the weak English feature value for verb-raising. Lardiere argues that these results show no syntactic deficit, but rather a difficulty with mapping fully specified syntactic features to surface morphophonology.

Goad and White (2006) also in an attempt to account for the variable performance on past tense by adult Chinese learners tested 10 intermediate Mandarin speakers on tense and participial morphology. The authors argued that the two languages differ with respect to the prosodic structure in the two languages; functional morphology is organized PWd-internally in Mandarin but it is adjoined to the PWd in English. Participants were tested on a combined sentence completion and production task. Participants did very well in both tasks; in production there were no significant differences between the suppliance of past (around 90%) and perfective (around 94%) or between regular and irregular verbs. Results from this study regarding Chinese speakers' performance on verbal morphology are not consistent with previous studies by Hawkins & Liszka (2003), Goad et al. (2003), Lardiere (1998) which by and large find a rather low performance of Mandarin speakers on inflectional morphology. As the researchers themselves argue it might have been a task effect as the design of the test could draw learners' attention to the structures under investigation.

More recently, Cabrelli Amaro et al. (2017) investigated the acquisition of past tense examining Spanish, Japanese, and Mandarin Chinese advanced learners of English. Spanish and Japanese have T specified for [past] while Chinese does not have a syntactic past feature. In addition, all three languages have similar phonological restrictions on consonant cluster formation required for the formation of past tense in English. A written sentence completion task and an oral sentence completion task were used. Results showed that in the written task, performance was very high for all groups (>93%). In the oral task though, none of the three groups had similar performance to native speakers. Therefore, having the [past] feature in their L1s does not entail target like performance. However, L1 phonology could not fully explain the results in the oral production; Chinese participants did worse (71%) than Spanish (77%) and Japanese (83%) although all three languages share the same phonological restrictions. This work casts doubt to the claim phonology to be the cause of variation in performance on past tense marking.

The emerging generalisation is that aL2 acquisition of agreement and tense morphology is quite difficult for Chinese learners even at advanced stages (cf. Table 4.1) whereas learners from L1 backgrounds where tense and agreement features are instantiated do better (e.g. Japanese, German, Spanish) although not necessarily at ceiling. It is an open question currently whether the challenge Chinese learners face is due to a syntactic deficit or mapping syntactic features to morphemes, potentially due to L1 related prosodic constraints.

Table 4.1: Results on 3SG -s and Past of L1 Chinese advanced L2 adult studies

	3SG -s	Past regular	Past irregular
Lardiere, 1998a, b	4.5%	34.5%	-
Hawkins & Liszka, 2003	-	63%	84%
Goad & White, 2006	-	90%	93%
Cabrelli Amaro et al., 2017	-	71%	-

Considering commission errors, they are not uncommon in adult L2 acquisition of verbal morphology; for example, finite forms being used in non-finite positions (Meisel, 1991 for German and French) or past tense verbs used in present tense contexts (Zobl, 1998; for Russian L1 learners of English) are not as rare as in L1 acquisition. L2 learners produce more commission errors than L1 children. L2 learners such as speakers of New Englishes usually overgeneralise the progressive aspect in habitual contexts as in the example '*Are you wanting a cup of coffee?*' (Schubert, 2002:1) while they also sometimes use stative verbs in progressive forms (Andersen and Shirai, 1996).[11] Al-Hamad et al. (2002) also found that advanced Japanese learners of English differed from native speakers in accepting the progressive tense when the context required a habitual reading while advanced Chinese learners of English in the same study did not reject present progressive with statives.

Another commission error that has been found in adult L2 learners' speech is the overgeneration of *be* in English as in example '*I was have a breakthrough*' (Lardiere, 2007:92) as well as in other Germanic languages. Such errors have been reported for learners with different L1s; Chinese (Lardiere, 2007: 92), Punjabi (Huebner, 1989) and different L2s e.g. L1 Turkish, Dutch L2 (Van de Craats, 2009). Interestingly, both untutored and instructed adult learners produce such forms. In the Dutch part of the ESF corpus, '*is + verb(-x)*' instances are very frequent and appear after a first stage which lasts 9 months when learners use only

[11] Input from British dialects may play a role or transfer from Indian but a general tendency by learners to overuse the progressive form is also recognised as a separate factor to account for the overuse of progressive forms.

infinitival forms. Huang & Yang (1998) also found the overgeneration of *be* in essay writings by university students in Hong Kong who were intermediate to advanced EFL learners. Finally, Parodi (2019) argues that '*is'* is used in L2 Dutch and German as well as other non-thematic verbs unproblematically to mark agreement. According to Parodi, these non-thematic verbs are finiteness markers, instances of T.

A final typical characteristic of adult L2 acquisition is the intra- and inter-individual variation (Schlyter & Thomas, 2012). The same learner may produce various target and target deviant forms while adult L2 learners as a group exhibits great individual variation.

To summarize, adult L2 acquisition of verbal morphology is characterized by persistent omissions of inflections as well as by a number of commission errors even at advanced stages of acquisition. In this respect, it differs from L1. There is evidence that aL2 use non-thematic verbs, in particular *be*, to mark finiteness. A similar phenomenon is observed in cL1 but appears more restricted and seems to be used mainly for 3SG agreement.

4.3 The acquisition of finiteness in English by L2 children

In this section, I turn to cL2 and compare immersed and instructed learners. I also review work on the age of onset which is central to this thesis.

4.3.1 The acquisition of finiteness in English by L2 children in immersion contexts: age of onset effects

Li (2012) investigated the acquisition of tense and agreement morphology in English by L1 Chinese children learning English naturalistically in their early stages of language acquisition. Six children aged from 7 to 9 years old with length of residence in the U.S. between 4 and 6 months were tested at different times over a seven-month period. Data were collected through an elicitation task based on picture description as well as through a general (free) conversation with the experimenter. Production of 3SG *-s*, past regular *-ed*, copula and auxiliary *be* as well as syntactic structures (such as overt subjects and nominative case assignment on subject pronouns) were examined. Copula *be* was consistently supplied by all children to a very high percentage (93%). Auxiliary *be* in progressive contexts was supplied in a lower rate but the author does not provide an average percentage of use for auxiliary *be*. Errors with respect to auxiliary *be* were considered separately; overgeneralization of auxiliary *be* preceding stem verbs in non-present progressive contexts (e.g. '*Mr. Darr every day is give everybody a lot of homework.*' 2012:84) was the most common error type followed by the use of the structure of '(*inflected) be + inflected verb'* as in '*They are went to forest'* or '*One dog was came'*

(2012:85). (These errors constituted between 2-72 instances in children's speech. Unfortunately, no percentages of these types of errors are provided out of the total number of verbs.) Suppliance of 3SG *-s* in habitual contexts was very low with omissions of the inflectional morpheme to be the most common error type. The average correct supply of 3SG *-s* in obligatory contexts in the recordings of all children across the testing sessions was 16%. Similarly, production of regular past *-ed* was quite low (average 13%) in contrast to the irregular past tense (average 38%) which was used somewhat more productively. Regarding commission errors, there were a few instances of verb inflection with *-s* when the subject was not 3SG; tense and number agreement errors were mainly attested with copula and auxiliary *be* (e.g. use of '*is'* instead of '*was'* in past context or use of '*is'* instead of '*are'*). Syntactic properties were in place with children's performance on overt subjects and nominative case assignment at ceiling (100%). The author concludes that affixal inflection requires more time to acquire than suppletive morphology and that in this early stage of language development tense and agreement morphology was not used productively by children while the related syntactic properties did not show to pose any problem to learners. To account for her findings, she proposes that abstract functional categories are present in the early stages of L2 grammar and syntax triggers the acquisition of morphology. This implies that no maturational constraints are assumed as learners are considered to have access to the abstract categories. However, Li does not explain in detail why children overgenerate *be*, a frequent pattern in her data; all six children produced such instances while for some of them there are quite a lot of instances.

Paradis et al. (2008) focusing on the early stages of the L2 acquisition examined the acquisition of tense in English by three groups of child learners, TDL1, SLI, and L2. TEGI (Test of Early Grammatical Impairment) was used to document production and comprehension of inflectional morphology (3SG *-s*, past *-ed*) and unbound tense morphemes such as the copula *be*, auxiliary *be*, and *do*-support. Twenty-four L2 children from various L1 backgrounds had mean age 5;7 and a mean of 9.5 months of exposure to English. Eight of them had a Mandarin or Cantonese background, languages which do not mark tense grammatically (Lin, 2001; Matthews & Yip, 1994). L2 learners performed quite low in inflectional morphology (3SG -s: 16% and Past -ed: 20%) and lower than L1 children (3SG -s: 42%, Past -ed: 47%) while the groups did not differ with respect to the production of unbound morphemes. They also tested subject-auxiliary inversion in questions and found similar performance for all groups. They concluded that their findings for the L2 children support the MSIH; they show optional use of inflectional morphology but do not have issues with syntax indicating unimpaired syntactic representations.

Ionin & Wexler (2002) tested 20 Russian child learners of English focusing on verbal morphology: copula and auxiliary *be*, third person singular -*s*, and past tense -*ed*. Children were aged between 3;9 and 13;10 with mean age being 8;4 and were all residents of the USA at the time of testing with varying lengths of exposure which would be less than three years. Results of spontaneous production data showed a high number of omissions of inflection especially for 3SG and past tense (78% and 58% respectively); omissions of inflection on '*be'* forms were much less (cop.:16%, aux.: 33%) showing an asymmetry between *be* forms and affixal morphology. Children made only a small number of tense/agreement errors while overgeneration of *'be'* mainly preceding stem verbs as in '*They are help people when people in trouble*' (2002:111) was used with a range of meanings: progressive, generic, stative, past, future, indicating that this use is not an incorrect present progressive form. Apart from these morphological properties, the authors also examined the verb placement of thematic verbs and auxiliaries with respect to adverbs and negation and concluded that children know their differential position. Based on these findings, Ionin & Wexler argue that children have the relevant functional categories of Tense and Agreement intact as they do not make tense/agreement errors, they are mostly accurate with forms of '*be'* and also know the verb placement of finite forms. They analyse the overgeneration of '*be'* as a substitute of affixal inflection in an attempt to mark tense and/or agreement on the verb. The authors argue that their results are in line with the MSIH as functional categories or features are fully in place and any difficulties with affixal morphemes are due to retrieval and communication pressures. The MSIH though cannot account for the asymmetry between *be* forms and affixal inflection. To explain this asymmetry, the authors suggest that the low inflection on thematic verbs may be due to a generalization L2 children make by associating morphological agreement to verb-raising. They propose that children initially have access to universal rules before they acquire the language specific ones. Denoting affixal inflection on unraised thematic verbs is suggested to be an English-specific rule which probably requires time for mastering, while expressing morphologically overtly checked agreement features is a UG requirement.

In addition to production data, Ionin & Wexler also administered a GJT to test children's comprehension of verbal morphology and although I will not go into detail here, it is worth mentioning four findings concerning their less advanced learners: 1) a high rate of acceptance (44%) of inappropriately inflected items (e.g. *I goes*), 2) 40% acceptance rate of items with omitted '*be'* forms, 3) 20% acceptance of incorrectly agreeing auxiliary forms and 4) the 19% acceptance rate of dropped -*ing* items, that is 'be + verb-∅'. The authors argue that these results may be explained by the nature of the test which might place processing and retrieval difficulties.

The picture emerging from the above studies is that cL2 learners show very low accuracy in verbal morphology, very high omission rates, and some commission errors (e.g. *be + verb(-x)*). However, based on syntactic properties researchers tend to assume that interlanguage representations are intact and features fully accessible. Let us now consider studies focusing on finiteness in English after several years of exposure.

Jia & Fuse (2007) investigated the acquisition of English verbal morphology including 3SG -*s*, regular and irregular past by immigrant Chinese L1 children during 5 years of immersion in the US. 10 children and adolescents with age of arrival ranging from 5 to 16 years of age formed two groups: a group of early arrivals (6 children arriving before the age of 9) and a group of late arrivals (4 children arriving after the age of 12y.o.). Results of spontaneous speech data showed considerable variation regarding the individuals in each group which achieved mastery, set at 80% accuracy in mastery of morphemes after 5 years of immersion. Specifically, considering only the 3SG and the past tense, only 3/10 children mastered the 3SG -*s* by the end of the study, 4/10 children mastered past irregular and no child mastered past regular. Carrying out a growth curve analysis, they found that it was the language environment composite score rather than AoAr (Age of Arrival) that better explained participants' performance. They argued that some grammatical morphemes (e.g. 3SG -*s* and Past) are harder to learn than others (e.g. progressive aspect, copula *be*) and require several years to master. It is quite interesting that so few children mastered the tense morphemes after 5 years of immersion. Although age of onset effects could not explain these findings as well as the language environment composite score did, one point worth mentioning here is that if age of onset effects occur and affect the acquisition of inflectional morphology earlier than the age of 5/6 (e.g. Meisel, 2009) then all children of this study are exposed to English at an older age (2 only children had some previous classroom instruction in English in their countries) meaning that we could perhaps indeed expect other factors to override the age of arrival effect.

More recently, Paradis et al. (2016) also examined age of onset effects investigating the acquisition of English L2 verb morphology by Chinese speaking children residing in Edmonton, Canada. The 18 participants in the study had mean AoO 4;2 and the range of AoO was between 1;7-5;8. Mean age at time of testing was 8;5 in round 1, 9;5 in round 2, and 10;5 in round 3. TEGI (Test of Early Grammatical Impairment) was used to document production and grammaticality judgments on 3SG -*s* and Past tense (among others). Considering the production data, significant change in children's performance was attested from round 1 to round 2 but not from round 2 to round 3. In round 3 which took place after six years of exposure, 13/18 children reached monolingual standards on 3SG -s and 15/18 children had acquired Past tense. A lot of individual variation was found in the children's long-term outcomes with L2 verb morphology. Regression analyses showed that performance on 3SG -*s* was significantly predicted by vocabulary size, English richness (outside school), and allomorph type (-s, -z, -iz); for Past regular: significant predictors were the verbal short-term memory, vocabulary size, and allomorph type (-t, -d, -id), while for Past irregular: verbal short-term memory, vocabulary size, and word frequency. Overall, subject-verb agreement marking was found particularly problematic for this population even after 6 years of immersion whereas acquisition of *be* was unproblematic. As far as the children's developmental trajectory is concerned, in line with Jia & Fuse discussed above, this study also reports that children reached a plateau after five years of exposure and not all children reached monolingual standards. Paradis et al. suggest that this is due to crosslinguistic influence from children's L1 Chinese which is an isolating language. More time is required for acquisition of verbal morphology for populations of isolating language backgrounds while fossilization may be the long-term outcome for some children.

Although age of onset effects were not found in this study either, we need to keep in mind that differences in the ages of onset among the children may not have been significantly different.

The picture that has emerged so far shows very high omission rates of inflectional verbal morphology in early stages - as illustrated on Table 4.2 below - while L2 children's performance is further characterised by commission errors (e.g. overgeneration of *be* attested by Li, 2012; Ionin & Wexler, 2002) in production but also in comprehension. These facts confirm the difficulties all learners have with verb morphology in English during the early stages. Through continuous exposure to input and after several years of exposure some children may acquire the morphemes while others may reach a plateau. However, we saw that for Chinese speakers learning English naturalistically problems seem to persist even after many years of exposure.

Table 4.2: Results on 3SG -s and Past tense by child L2 learners of English

	Length of exposure	L1	3SG -s	Past regular	Past irregular
Paradis et al., 2008	9.5 months	Various L1s	16%	20%	-
Li, 2012	(≈)1 year	Chinese	16%	13%	38%
Ionin & Wexler, 2002	> 3 years	Russian	22%	42%	58%
Jia & Fuse, 2007	5 years	Chinese	44.82%	40.26%	65.65%
Paradis et al., 2016[12]	6 years	Chinese	96%	97%	85%

With respect to age of onset effects, all of the above discussed studies claim that other factors could better account for children's performance such as the language environment. They further argue that L2 children have not impaired representations with features being accessible to them. Not all studies agree on that though. I now turn to some studies indicatively considering language pairs other than English and other than finiteness phenomena.

Armon-Lotem et al. (2011) tested Russian-speaking children with either Hebrew or German as an L2. Children's age of onset was average 2.4 for Russian-German and 2.9 for Russian-Hebrew and age at time of testing was on average 5.5 and 5.10 respectively. Children were divided into three groups according to their age of onset and their performance was compared. Results were slightly different between the two cohorts; for the Russian-Hebrew learners, they found that age of L2 onset correlated negatively with L2 grammatical performance (meaning that the earlier the better) while this was not found for the Russian-German speakers. The reason for this as they claim might be the very narrow age of onset range, i.e. R=12-46 for the Russian-German and R=0-66 for the Russian-Hebrew children.

Age of onset effects were also reported by Kroffke and Rothweiler (2006) and Rothweiler (2006) who found that Turkish-speaking children with AoO=3-4 patterned as monolinguals with respect to word order, subject-verb agreement and subordinate clauses in German, while those children with an AoO> 4 patterned similarly to L2 adults.

[12] Considering the differences in percentages of suppliance of 3SG-agreement -s and past tense by subjects in Jia & Fuse (2007) and Paradis et al. (2016) studies after similar number of years of exposure, it seems that the difference may lie in the age of onset of subjects. Jia & Fuse's subjects had AoO that ranged between 5 and 16, while in Paradis et al.'s study, children had AoO that ranged between 1;7- 5;8.

Finally, Meisel (2008a) examining German L1 children learning French naturalistically (in an institutional setting: day school) found that 6 out of the 10 children patterned with L2 adults in the acquisition of finiteness while the rest 4 exhibited performance similar to (2)L1s, a fact that was explained as the impact of the AoO.

Overgeneration of be

As we saw, when discussing the acquisition of verbal morphology, an error pattern attested in several studies (e.g. Li, 2012; Paradis et al. 2008; Ionin & Wexler, 2002) is the '*be + verb'* structure. In fact, it has been found in early stage learners' interlanguage regardless of their L1: in Russian learners (Ionin & Wexler, 2002), in Chinese learners (Li, 2012; Yang and Huang, 2004), in Spanish learners (García Mayo et al., 2005), in learners of various languages (Paradis et al. 2008), but also in other L2s e.g. Dutch and German (Van de Craats, 2009). Its use is not restricted to progressive contexts, on the contrary, it is used with various meanings, thus we talk about *be*-overgeneration. It has mainly been found as '*be + stem verb*', however, '*be + verb-ing*' has also been used with a range of (non-target) meanings (Hawkins & Casillas, 2008).

Let us consider some of the explanations proposed for this widespread phenomenon. Ionin & Wexler (2002) found that this phenomenon in Russian children's spontaneous speech accounted for a quarter of their utterances with an overt, finite *be* auxiliary (or 9% out of all inflected utterances). The structure appears with various meanings as in the example: '*they are help people when people in trouble*' [They help people when people are in trouble] (Ionin & Wexler, 2002: 111) where it is used to denote generic meaning. There are examples where the intended meaning would be past, future, stative, etc. The authors examine whether this could be an effect of transfer, however, they exclude this possibility as Russian lacks an overt copula *be* in the present tense as well as auxiliary *be* (except compound future tense). They argue then that the auxiliary *be* is used as a tense/agreement marker. L2 learners first acquire morphological agreement on *be* forms which are raised overtly to Tense while affixal inflection is acquired at a later stage as 'long-distance Agreement does not initially trigger morphological agreement for L2 learners' (ibid, 116). In other words, L2 learners know that *be* forms need to be morphologically realised but have not mastered the English language specific rules concerning morphological agreement on unraised thematic verbs. This is taken as proof that UG is available while parameter resetting requires time. A similar proposal is made by Parodi for adult L2 acquisition (see Section 4.2).

Paradis et al. (2008) also found this pattern in her studies of children with various L1 backgrounds including morphologically rich L1s (e.g. Spanish, Romanian, Ukrainian) as well as isolating L1s (e.g. Mandarin, Cantonese). Thus, transfer cannot account for its use. Although the authors do not totally exclude the proposal made by Ionin & Wexler (2002) that *be* may work as a finiteness marker in early grammars, they also discuss another possible explanation, namely the input frequency and distributional consistency. As *be* is more frequent and consistently distributed in the input, it might mean that it is stronger than inflected forms *(-s*, -*ed*) and thus its retrieval is easier.

More recently, Hawkins & Casillas (2008) accounted for the occurrence of this structure (i.e. '*be +verb'*) within the framework of the Contextual Complexity Hypothesis. Under this proposal, '*be* + *verb*' is a possibility, as forms of '*be'* refer to T and the properties of the subject, allowing for a bare verb to follow. However, learners are not expected to produce '*be + inflected verb'* (e.g. **she is played*) because they 'specify contexts of insertion for /s/ and /d/ early on in terms of a T that is itself specified for [+/- past], but not for [BE]' (Hawkins & Casillas, 2008: 603).

Finally, Gavruseva (2008) argues that the *be* overgeneration is a generic finiteness marker which is produced due to the misanalysis of the input. It is expected to occur with statives and punctuals as these are expected to be finite under her Underspecification of AspP Hypothesis. Most of the reviewed studies have not considered aspect and thus, we do not know if this is the case.

4.3.1.1 Summary

To sum up, the acquisition of inflectional morphology is problematic for all (early) learners. L2 children's interlanguage grammars with respect to the acquisition of finiteness has been found to be characterised by

- the higher use of copula *be* than auxiliary *be*
- the higher use of copula and auxiliary *be* than affixal morphology on thematic verbs
- the omission of inflections (especially 3SG -*s* and past -*ed* in English)
- the rare overgeneralisation of inflections (commission errors)
- the overgeneralization of the constructions '*be* + *verb-ing', 'be* + *bare verb',* and *'v-ing'* with various meanings

Continuous exposure to input may lead to acquisition but this may depend as well on the learners' L1. For Chinese learners, for instance, many years may be required before mastery while fossilization cannot be excluded as a possibility even for a learner in an immersion context including L2 children.

The impact of age is an unsettled issue in these studies as in some studies age is a predictor while in others input/language environment are better predictors of the data. In addition, there is no consistency with respect to the factors proposed to account for the L1 – L2 differences. Thus, some studies propose the age of onset, others the amount of exposure, and yet others the working memory, the vocabulary size, etc. Finally, there are methodological issues that need to be tackled such as the narrow age spans, the lack of control of confounding variables such as chronological age or length of exposure, and the small samples. Most studies have included only a small number of children who are further subdivided in age groups making the numbers even smaller and any comparisons between them cannot provide robust evidence.

Undoubtedly, this picture of the age effects within childhood as well as the fact that there is only a limited number of such studies call for further research. In order to be able to uncover such effects, we need to consider a wider age window and a much bigger sample that could let us if not pinpoint a cut-off point, reveal the gradual nature of the decline in performance if this exists.

4.3.2 The acquisition of finiteness in English by L2 children in instructed settings: age of onset effects

The role of age in child SLA research has attracted researchers' interest for different reasons; more theoretically-driven researchers aim to shed light on whether maturational effects exist and how these impact on the human ability to learn languages; others are concerned with practical questions about children's learning in instructed contexts to inform practice and policy.

There are a few studies -to the best of my knowledge- to investigate the acquisition of finiteness in English by L2 children in a classroom setting. Let us start with them.

Garcia Mayo et al. (2005) tested 3 groups of bilingual Basque/Spanish children learning English in school after 4 years of exposure. The first group of younger children (N=20) were 7-8y.o. at time of testing and started learning English at age 4-5; the second group (N=20) were 12-13y.o. at time of testing with age of first exposure at 8-9; and the third group (N=18) were 14-15y.o. with age of first exposure at 11-12. They elicited oral narratives and found differences between the groups: the younger group used independent lexical verbs (i.e. not accompanied by '*is*') to a very limited extent (27.33%) and relied mostly on *'is'* default forms (use of '*is'* even in plural contexts), older groups showed a different pattern using independent lexical verbs to a high degree and default forms were limited. Considering inflected forms, these were only examined for the older groups and accuracy rates were quite low; Group 2:

15.12%, Group 3: 43.15%. Younger learners' performance was also characterised by the overgeneration of '*is'* preceding a lexical verb. It accompanies both transitive and intransitive verbs, as well as bare and inflected verbs. It is also noteworthy that it appears as '*is'*, as in the examples: (1) a. *the kid is open the door*, b. *the boy is came*, c. *the little boy is want the frog*, d. *the boy and the reindeer is run*, e. *David and the dog is see two frog*, f. *the boy and the dog is sit down, g. the frog went out of the bottle and is escape from the house* (Garcia Mayo et al., 2005:447, 466). Finally, it is used with a range of meanings – the majority of which are generic as the authors claim. Interestingly, the younger group produce such forms to the greatest extent. The percentage of usage decreases with increasing age of onset of the groups. The oldest children – that is, those with the latest age of onset- do not use almost at all this structure. The authors interpret these results as insensitivity to inflection by younger groups, suggesting that parameter resetting has not taken place yet. As for the overgeneration of *'is'*, it is interpreted as an agreement morpheme which is transferred from the learners' L1 corresponding agreement morphemes and their positions associated with the verb.

Housen (2002) carried out a corpus analysis using oral data from the Corpus of Young Learner Interlanguage (CYLIL) to investigate the acquisition of verb morphology in English by Dutch and French L2 children. Learners were students of the European Schools in Brussels and Mol (Belgium) (N=46) and were grouped into four proficiency levels: a low (L), lower intermediate (LI), higher intermediate (HI) and high (H), each assumed to correspond to a different stage of L2 development. Ages of participants ranged from 9 to 17 with increasing age being associated to a higher proficiency level. Analyses of the L2 children's data showed that the vast majority (>80%) of the verbs produced by L and LI groups were base forms, i.e. uninflected, a percentage that decreases by increasing proficiency level. V-ing was the first inflected form to emerge, used by L and LI groups, but is randomly used in finite and non-finite contexts while it appears both with and without the auxiliary *be*. Data from HI and H groups show that its use becomes increasingly target-like being produced as a participle accompanying the auxiliary *be*. The next inflected form is the *Ven*[13] form with first verb emerging being *be* followed by *have* and *do*; they are rare in early proficiency levels but increase gradually. Next, regular *-ed* emerges which is not produced productively before the

[13] *Ven* is used as a short-hand notation for irregular Preterit or past participle-like forms (*went, seen, eaten*), (Housen, 2002:109).

HI level. Finally, the 3SG morpheme -*s* is rare or erratic in the two low levels, L and LI. Housen proposes three stages in the acquisition of verb morphology in English: Stage 1: 'Invariant default forms', Stage 2: 'Non-functional 'allomorphic' variation', during which forms and meanings are not mapped as in the target L2, Stage 3: 'Distributional restructuring, functional specification and increasingly target-like use of verb morphology to encode tense, aspect and agreement', that is, when forms and meanings are mapped together to a more target-like fashion. Not all learners but only a few from the H group reached this final stage. Housen concludes that the acquisition of verb morphology in English by instructed learners reveals its complex nature which is argued to be similar in route to that of older learners in naturalistic settings.

Yang & Huang (2004) investigated the acquisition of tense and aspect in English by Cantonese Chinese child, adolescent and adult L2 learners in classroom contexts. Specifically, 5 groups of learners were recruited for this study from primary and secondary schools as well as a university in Hong Kong: 1) 10-year-olds (N=270), 2) 12-year-olds (N=49), 3) 14-year-olds (N=56), 4) 16-year-olds (N=30), and 5) 19-year-olds (N=48). Learners attended 8-9 English lessons (i.e. mean= 4.5 hours) per week in primary school and 8-10 English lessons (i.e. mean= 5 hours) per week in secondary school. Note that English teaching starts at school from primary one, with the group of youngest children to be first tested at primary five when they are 10 years old. Each group had a different proficiency level ranging from late beginners (10-year-olds) to advanced (19-year-olds). Tense and aspect morphology were assessed through written narratives which were required to be narrated in past time. Results showed the following: 1) the group of the youngest children's finite forms were restricted to simple past tense and no other past tenses – this pattern gradually changes on the basis of proficiency with higher proficiency learners using a wide variety of structures in past, 2) the group of the youngest children showed a low usage of adverbials (9.7%); adverbial usage also gradually increases with proficiency; it stabilises around the range of 21- 31% as the university learners construct more complex and varied structures creating various past contexts, 3) the overall acceptable past marking increases with age/proficiency (44% - 44.2%- 56.7% - 67.5% - 82.2% respectively for each age group). It is noteworthy that for 6 years (10-16) learners produce past to a percentage up to 67.5%. Based on this finding, the authors argue that acquisition of the tense-aspect system in English does not find empirical support. Learners are said to have acquired the tense-aspect system of English only when they are already university students. Interestingly, the group of these advanced learners do not perform at ceiling (82.2%) either. Recall at this point that the methodology to assess production of tense-aspect in English was written narratives which can show their knowledge a bit inflated as learners have the opportunity to edit and correct their writings. Finally, they attribute low suppliance of past marking for 6 consecutive years to L1 influence. As Chinese does not mark tense grammatically, learners rely on pragmatic and contextual cues as well as lexical devices as they do in their L1 for an extended period of time.

Lee & Huang (2004) in what seems to be a follow up study, focused on *be* productions in the written narratives of the 9-10-year-old children as presented above (in Yang & Huang, 2004) which was their least proficient group. While copula *be* was said to be rather unproblematic (80% accuracy), auxiliary *be* proved to be quite difficult (10% accuracy). Copula *be* is acquired earlier than auxiliary *be*, a finding that is in line with previous literature. It is interesting though that copula *be* accuracy rates differ somewhat when considering the type of sentence. Thus, for example, while *be* preceding a noun shows an accuracy rate of 92% as in '*I am a king.*' (2004:216), copula *be* preceding a preposition as in '*The queen is in the palace.*' (2004:216) shows a respective rate of 62%. Accuracy on copula *be* then seems to depend to some extent to what follows it. Turning to auxiliary *be*, the main type of error was its usage with a main verb form, bare or inflected (262 occurrences of '*be + V'*/380 errors). Thus, a very high percentage of the errors involved examples like: a. **The queen is walked into her bedroom.* b. **They were came back*., c. **I am can make one clothes*., d. **He is open the door.*, e. **She was ask him*. (Lee & Huang, 2004: 218). The overgeneration of *be* is found mostly with activity verbs as well as verbs with a related adjectival form (e.g. open, die). Overall, it is argued to be a developmental feature which further reveals the creative process of L2 acquisition and shows the hypotheses learners make with respect to the use of *be*.

Apart from the studies presented above which scrutinised the acquisition of finiteness, there are some studies which are more policy-oriented and focused on examining macro-skills such as listening, reading, or grammar (e.g. Jaekel et al. 2017; Cenoz, 2003; Muñoz, 2006). Considering only those which focus on overall grammar abilities, I present a brief overview of such studies which usually also compare younger and older starters in learning an L2. The majority of these studies assess the rate of development and not the ultimate attainment of learners with the majority of them pointing to an older age advantage within childhood.

Studies such as the one by García Mayo (2003), the Barcelona Age Factor Project (Muñoz, 2006; 2010), the study by Myles & Mitchell (2012) have all tested L2 children with different ages of onset in classroom settings and compared their achievement on grammatical abilities -among others. The vast majority of these studies -if not all- show that older starters are more efficient learners than younger children while they further show that older learners have more developed cognitive maturity and employ a wide variety of learning strategies (e.g. Muñoz, 2006, 2010; Myles & Mitchell, 2012).

So far, we have been discussing L2 development of grammar considering both the route and the rate of development. Not many studies have looked into ultimate attainment of L2 grammar by instructed learners. One such study has been carried out by Larson-Hall (2008) who found some modest results in favour of earlier starters and argued that age does play a role in second language learning as long as enough input is provided. Pfenninger (2014), on the other hand, testing groups of learners with different ages of onset as well as exposure to English with respect to grammar did not find any significant difference between them. The author argues that perhaps the reason for this may be due to the fact that all groups practiced English grammar to the same extent despite the larger amount of exposure of some groups or the earlier age of onset. In any case, with respect to the ultimate attainment of grammar by classroom learners, no conclusions can be drawn for the time being due to the scarcity of data. Note that many studies have found earlier age advantages with respect to other skills such as listening comprehension but the puzzle with the grammar is still unsolved.

4.4 Comparison of L2 learners in immersion and instructed settings

Let us now turn to the comparison between immersed and instructed L2 children on the basis of route in the acquisition of finiteness, the rate of acquisition, as well as the ultimate attainment.

- Route

Starting with the route, that is, whether children in both settings differ qualitatively, it seems that L2 learners pattern the same in both contexts. In the early stages irrespective of the learning setting, L2 children show:

- higher accuracy on copula *be* > auxiliary *be* > affixal morphology
- very high rate of omissions of inflections
- not frequent overgeneralisation of inflections
- overgeneration of progressive aspect to other contexts
- overgeneration of the ‘*be + verb*’ structure

Qualitatively the performance of L2 children in both contexts and regardless of the input received seems to be the same. Quantitatively though there may be differences between them as immersed children may progress faster than instructed learners due to the abundant input they receive compared to the minimal input offered in instructed settings.

- Rate

As far as the rate of grammatical development is concerned, again this is similar in both situations in the sense that research in both naturalistic (e.g. Krashen et al. (1979; 1982) and classroom learners (e.g. Muñoz, 2006, 2010; Myles & Mitchell, 2012) has shown that older children progress faster than younger ones.

- Ultimate attainment

With respect to ultimate attainment, no definite conclusions can be drawn. In naturalistic settings, L2 children may end up native-like speakers of the L2 (e.g. in case of immigrant families whereas children became native-like while parents had persistent difficulties in the L2), however, as we saw earlier in this chapter some children may fossilise or reach a plateau despite living in the L2-speaking country. Research is scarce in ultimate attainment of instructed children; thus, further research is needed to clarify this issue. Table 4.3 summarises the overall picture as discussed so far.

Table 4.3: Immersed L2 children vs instructed L2 children with respect to morphosyntax

<table>
<tr><th></th><th>Route</th><th>Rate</th><th>Ultimate attainment</th></tr>
<tr><td>Naturalistic L2 children</td><td rowspan="2">NcL2 - IcL2: qualitatively similar</td><td rowspan="2">older children’s advantage</td><td>younger learners’ advantage</td></tr>
<tr><td>Instructed L2 children</td><td>not enough evidence yet</td></tr>
</table>

4.5 Interim Summary

Considering all the acquisition types discussed so far, there is agreement that L1 children differ from L2 adults in the acquisition of morphosyntax as exemplified by the acquisition of finiteness discussed in this chapter with researchers disagreeing as to which reasons underlie the differences in their interlanguage development.

What is still unknown is where child L2 acquisition lies with respect to the other two acquisition types; child L1 and adult L2. Research has shown that in some respects, child L2 resembles adult L2 (e.g. in terms of transfer) while in other respects it resembles the L1 child (e.g. in ultimate attainment).

Furthermore, it is still under investigation whether the route of early and late starters within childhood is the same and whether in both cases this is like child L1 acquisition or adult L2 acquisition in certain domains of grammar or whether groups with different ages of onset pattern differently as a function of decreasing age of onset.

Table 4.4 provides an overall picture of what research has shown so far and what we still need to find out.

Table 4.4: Children vs adolescents/adults with respect to morphosyntax

	Rate	**Ultimate attainment**	**Route**
Naturalistic settings	older L2 learners' advantage	younger learners' advantage	cL1 ≠ aL2 cL2 – aL2? cL2- cL1?
Classroom/Instructed settings	older L2 learners' advantage	?	L1 ≠ L2 cL2 – aL2? cL2- cL1?

Context does play a role but as we saw in the case of L2 children, there do not seem to exist qualitative differences because of the different learning modes; L2 children in both immersed and classroom contexts show similar error patterns as far as finiteness is concerned. In any case, we still need to establish that L2 children go through the same developmental stages in both settings as evidence is not yet robust. The same has been shown for adult L2 learners in immersed and instructed settings (e.g. Ellis, 1984a; Bardovi-Harlig 1992, 1995).

4.6 Studies investigating transfer effects

Let us, finally, consider previous studies dealing with transfer effects in child L2 acquisition. Haznedar (1997) investigated a Turkish boy, Erdem, in detail, reporting the existence of transfer effects in the word order of English in the first months of acquisition. Later, Whong-Barr and Schwartz (2002) tested L1 English, Japanese and Korean children in the acquisition of the English dative alternation using an oral GJT and results revealed significant differences between the two (Korean and Japanese) groups due to transfer effects in the double-accusative construction with for-datives where the two languages differ in relation to English. Similarly, Unsworth (2004) using a task that required both truth value judgment and production, compared English speaking children learners of Dutch to L1 children and L2 adults and found that cL2 learners resemble aL2 learners in the acquisition of direct object scrambling in Dutch, that is, they passed through the same developmental stages. In addition, transfer effects were observed in both cL2 and aL2 which were argued to stem from their L1 English which does not have scrambling. All these studies lend support to the FT/FA model discussed in Chapter 3.

There are also studies that found limited L1 influence in child L2 acquisition. Zdorenko and Paradis (2008) asked whether there are transfer effects in the acquisition of the article system in English by 17 L2 children from [+article] and [-article] backgrounds. The study was longitudinal and lasted for two years and at the first round children had a mean age of 5;4 and mean exposure of 9 months. Results revealed that the group with the [+article] languages did not transfer definiteness from their L1 (perhaps due to mapping problems) while children from [-article] backgrounds omitted articles more often than the group of [+article] languages but only in the very beginning of their acquisition. They thus argue that children go through a fluctuation stage and L1 transfer is limited in child L2 acquisition. Similar results were replicated in another study by Zdorenko and Paradis (2011) in the case of children with less exposure to the L2 (between 2-18 months).

Yet other studies found no L1 transfer effects. Paradis et al. (2008) examining L2 children from various L1 backgrounds on the acquisition of tense morphology in English at the early stages found no differences between children whose L1s were tensed and those whose L1s were tenseless (see section 4.3.1 for more details on this study).

In sum, the majority of the studies have reported some level of transfer effects in child L2 acquisition of grammar. However, it is not clear yet whether these transfer effects are only short-term or not. Furthermore, the domains investigated so far are diverse not allowing us to pinpoint the level of the language (e.g. morphology, syntax, interfaces) that transfer may occur. Then, the fact that the results with respect to the existence of transfer do not fully converge might have to do with the small samples some studies used as well as with the different phenomena explored. Finally, the relation between transfer and age has not yet been defined. Is transfer stronger after a certain age? Indeed, some studies seem to point to that direction (e.g. Blom & Bosma, 2016). Unsworth et al. (2014), for instance, suggested that transfer effects in the acquisition of the Greek grammatical gender become stronger over time as English/Greek bilingual children with a later AoO (and thus more entrenchment in a language without grammatical gender) performed worse than children with an earlier AoO.

A final note on age of onset effects and transfer is important. Transfer effects in early bilingual children have been reported to partly depend on the language combination under investigation (Unsworth et al. 2014). That is, the extent of L1 transfer will be affected by the typological similarity/distance of the languages examined as well as of the AoO of acquisition. It has been suggested that the larger the typological distance between two languages, the more persistent the difficulties the L2 children have might be (Blom & Paradis, 2016). In fact, it has been shown that L2 children from isolating L1 backgrounds require more time to acquire morphological properties than L2 children from inflecting L1s who omit inflections to a smaller degree (Chondrogianni, 2018). It would then be interesting to contrast children with different ages of onset and from different language backgrounds in the same grammatical phenomena in order to elucidate this issue.

Following this discussion of previous studies with respect to the impact of age of onset and L1 influence on the acquisition of morphology by various types of learners and having highlighted the open issues in literature, the next chapter presents the current study with the research questions and hypotheses aiming to address these gaps in literature.

CHAPTER 5

THE PRESENT STUDY: RESEARCH QUESTIONS, HYPOTHESES AND METHODOLOGY

5.0 Introduction

The impact of age of onset on cL2 remains unclear in many respects partly due to conflicting empirical results, as discussed in the previous chapter. The present study aims to investigate age effects on the acquisition of verbal morphology in English in conjunction with the effects of the learners' L1. In section 5.1, I present the research questions and hypotheses and in section 5.2 the methodology followed.

5.1 Research Questions and Hypotheses

Research Question 1: What is the impact of age of onset on the acquisition of tense and agreement features in English by L2 children? Specifically, will cL2 pattern with adult L2 (as hypothesised by Meisel, 2009) or with cL1 (as predicted by e.g. Schwartz, 2004)?

Hypothesis 1: We expect inflectional morphology to be a challenging domain for L2 children who will pattern qualitatively similar to L2 adults in this respect rather than L1 children. The reason I adopt this hypothesis is that empirical findings as discussed in detail in the previous chapter point to considerable problems with inflectional morphology by L2 children (e.g. Ionin & Wexler, 2002; Li, 2012).

Prediction 1: We thus predict omission errors, but also error patterns/commission errors attested only in aL2 and not in L1 acquisition.

Research Question 2: Is there a gradual decline in L2 children's performance in acquiring inflectional morphology depending on different ages of onset? Will early learners outperform later learners in the acquisition of functional features of T?

Hypothesis 2: Early learners will do better in inflectional morphology than later learners. Theoretically, if age 4 marks the end of the optimal period, children starting at 4 may have an advantage being closer to this optimal window than later learners. Further, many empirical studies, e.g. work by Johnson & Newport (1989) showed that there is a gradual decline in performance for later starters (which flattens in adulthood).

Prediction 2: Early starters are predicted to be more accurate in verbal morphology in English than later starters.

Research Question 3: Is there L1 influence? As reviewed in previous chapters, the features of [past] and [agreement] are not selected in Chinese as they are optional (Hawkins & Liszka, 2003) while they exist in Russian. Will Russians outperform Chinese?

Hypothesis 3.1: Following the Full Transfer/Full Access approach, L1 effects will be observed at the lowest proficiency learners because of positive transfer and L1 effects disappear at more advanced levels, because more proficient learners will have acquired the features.

Prediction 3.1: Russians will produce the relevant morphemes earlier than Chinese learners at the early stages. Both are expected to perform similarly at later stages as an effect of increasing proficiency.

Hypothesis 3.2: According to the Representational Deficit approaches, L1 effects will progressively be visible with increasing proficiency; learners who cannot rely on their L1 for the relevant features will have difficulty acquiring them.

Prediction 3.2: L1 effects are particularly expected at later stages, Russians are predicted to outperform Chinese learners who will show persistent difficulties.

Research Question 4: Are some inflectional markings more challenging than others?

Hypothesis 4.1: Interpretable features will be less problematic than uninterpretable features (following the Interpretability Hypothesis).

Prediction 4.2: Past tense will be less problematic than 3SG-agreement.

A final issue we will attempt to address if the answer to RQ1 is that cL2 is like adult L2 rather than cL1 (which is our hypothesis), is why cL2 is different from cL1 and what causes the difficulties with verbal morphology. Are there defective features (Representational Deficit approaches) or learners' performance is a question of mapping (e.g. MSIH)?

Let us now turn to the methodology followed to address these questions and hypotheses.

5.2 Method

I begin the presentation of the methodology with a description of the instructional setting in EF schools and background educational context in Shanghai and Moscow. I then discuss the design rationale followed by a description of the subjects, the tasks used, and the data collection process. I also consider the issues of length of exposure, input and proficiency in separate sections.

5.2.1 Instructional setting in EF

This research was carried out in EF (Education First or English First) schools in Shanghai and Moscow. EF English First Schools are schools teaching English as a Foreign Language to children 3-17.

There are many such schools around the world following a similar curriculum but with some variation in number of teaching hours offered across countries. I here focus on practices in China (Shanghai) and Russia (Moscow). Table 5.1 presents the academic hours offered per week in EF schools in the two countries.

Table 5.1: Academic hours offered for the various levels and ages by EF in China and Russia. (Note that one academic hour lasts 40 minutes.)

Learners' ages	**China**	**Russia**
3 - 6	3 academic hours per week	3 academic hours per week
7 - 9	3 academic hours per week	4 academic hours per week
9 - 17	3 academic hours per week	5 academic hours per week

EF employs native speakers of English or qualified non-native speakers and provides variable teacher training. In Shanghai, teachers are mostly employed native speakers of English while in Russia most teachers are Russian native speakers. All teachers follow the same EF internal curriculum and make use of the same materials. EF creates their own book series as well as online/interactive activities which are used in their schools worldwide.

Teaching follows mostly the communicative approach blended with some behaviourist elements (e.g. drilling, rewards) especially for younger learners. With older learners, teaching focuses more on form than for younger learners although the approach is again communicative, but drilling is abandoned.

Thus, the settings in the two countries are fairly similar with the only exception being the fact that in China there are more native speakers teaching than in Russia.

5.2.1.1. EF books and introduction of grammatical phenomena

Children starting learning English in EF at age 3 follow the Small Stars book series up to the age of 6. The Small Stars series covers the A0 CEFR level and only basic structures are introduced. From age 7 to 9, the High Fliers books are used which take children up to the A2 CEFR level.

A child starting at EF at age 4 will be exposed to Simple Present and Present Continuous early on with the former expected to be mastered at the ages of 7-8. Simple Past is introduced later than the Present Simple at the ages 7-8 (High Fliers F book) to be mastered at the ages of 8-9. Table 5.2 below presents the above information schematically.

Table 5.2: Information about the book series[14] and the time the phenomena under investigation are introduced

	Small Stars	**High Fliers**	**Frontrunner**
No of books	A - C	C - J	1-16
CEFR level	A0	A0 - A2	A1-C1.2
Children's age	3-6	7-9	9-17
Age of onset of a child learning English at EF	3	7	9
Features and structures	- Pres. Simple - Pres. Continuous	- Pres. Simple - Pres. Continuous - Simple Past	

We can already see that there is an important difference in the curriculum of the younger and older children in that children who attend EF classes from 3 to 6 years of age have a prolonged A0 CEFR level.

5.2.2 Educational context in China and Russia

Since our study involves cross-national comparisons, I briefly consider the general educational context as it may play a role in children's performance. With respect to L1 literacy, in Shanghai,

[14] Discussion of the material and the curriculum is based on the series being used at the time the research took place.

literacy officially starts at age 6 (primary one) in day schools. Preschool education starts at age 3 up to 6 and is not obligatory although the enrolment rate in urban areas is very high (98-99%) (Zhou, 2011). In preschool education, children are not explicitly taught academic skills but teachers follow informal literacy practices (China Preschool Education Research Society, 1999 as referenced in Li, 2013; Li et al., 2008). In Moscow, L1 literacy also starts when children enter state day schools around the ages of 6.5-8 (Bodrova & Yudina, 2018). Preschool education is not compulsory but similarly to Shanghai the vast majority of parents enrol their children to preschool education whose goals are to take care of the children and offer educational services (Bodrova & Yudina, 2018). The age of onset of L1 literacy then seems to be slightly earlier in Shanghai, something we also attested when considering answers to relevant questions in parental questionnaires. Chinese parents answered that on average their child started to learn to read and write in Chinese at age 4 while in Russia the respective average answer was 5.

There are further differences in cultural attitude towards schooling. In Shanghai, there is a significant tendency towards early learning and a general push on children to achieve a lot early on. As Sun et al. (2014) argue Chinese parent's mindset is along the lines of "the earlier, the better" and "we must not lose at the starting line" (2014; 2). Education is of the outmost importance and a top priority for Chinese parents who have very high demands and expectations from their children (Hu & Szente, 2009). Further considering the 'one-child policy' in China held until recently, parents paid even greater attention to their child's academic achievements (Hu & Szente, 2009) as it is seen as the determinant for their social upgrade in a very competitive market. Parents become anxious about their children's progress already when they reach preschool age (Hu & Szente, 2009). It is not surprising then that currently in China private English language schools but also early learning centres aiming to boost children's cognitive and general learning development are proliferating. Children in EF in China start early but around the age of 15 they quit so they can prepare for the 'Gaokao' exam, which is the national exam that will determine their entrance to university.

In Moscow, children generally start English later (not only privately but in day schools as well) and continue attending classes at EF up to the age of 17 or so. Unfortunately, there is no much literature on parents' attitudes towards education (at least in English language). However, based on personal experience and personal communication with Russian people in Moscow during the data collection, it seems that Russian parents do also value education but are not as anxious as Chinese parents early on.

5.2.3 Study design & rationale

This study's participants were (TD, i.e. Typically Developing) Chinese and Russian children learners of English. As inclusionary criteria, participants should be either 9 or 12 years old at time of testing and should have had 5 years of instruction at EF. I wanted to ensure that 5 years would guarantee sufficient exposure for children to do the tasks and narrate stories. In addition, these criteria aimed to control for the biological age at either 9 or 12, age of onset in learning English at 4 and 7 respectively and thus the same or at least comparable amount of exposure of learning English at EF.

Table 5.3 presents the design of the study with its criteria. In what follows, I discuss each L1 cohort in turn.

Table 5.3: Study design and criteria for inclusion applied to both Chinese and Russian learners

PARTICIPANTS	**Age of onset (AoO)**	**Age at time of testing (AT)**	**Length of exposure**
Chinese & Russian children	4	9	5 years
	7	12	5 years

5.2.4 Subjects

5.2.4.1 Chinese children

A total of 73 child participants were tested during the data collection which took place in various EF schools at Shanghai, China. Table 5.4 summarises information of participants. All children were native speakers of Chinese and were learning English as a foreign language (EFL). Participants formed two groups according to their age at time of testing and age of onset in English; the first group involved 39 9-year-olds (20 girls and 19 boys) and the second group involved 34 12-year-olds (16 girls and 18 boys). The younger group had a mean biological age of 115 months or 9;7, while the older group had a mean age of 150 months or 12;6.

Table 5.4: Chinese subjects

	No of participants	**Mean Age**	**Range**
9-year-olds	39 (20F, 19M)	9;7	9;0 – 10;1
12-year-olds	34 (16F, 18M)	12;6	12;0 – 13;9

Children were not exposed to English at home, as the home language as well as the parents' L1 was always Chinese including Mandarin or Cantonese with a few also reporting Shanghainese which is typologically similar to the Mandarin and Cantonese dialects.

Children come from families of a high socioeconomic status with highly educated parents. As seen in Table 5.5, the vast majority of the parents (124/144) have a university degree (>85%); only a few (20/144) have only professional training or low/upper secondary education.

Table 5.5: Overview table of background variables extracted from the parental questionnaires of Chinese children

	9-year-olds	**12-year-olds**
Cultural capital	Shanghai	Shanghai
Parents' L1	Chinese	Chinese
Home Language	Chinese	Chinese
Parental education	No degree: 5/76 (6.5%) Undergraduates: 50/76 (65.5%) Postgraduates: 21/76 (27.5%)	No degree: 15/68 (22%) Undergraduates: 36/68 (53%) Postgraduates: 17/68 (25%)

5.2.4.2 Russian children

A total of 74 children participated in this research who were learners of EF schools in Moscow, Russia. Table 5.6 summarises information on participants. The same inclusionary criteria were applied: 9-year-olds and 12-year-olds with age of onset at 4 and 7 respectively. As with the Chinese children, two groups were formed, 32 9-year-olds (14 girls and 18 boys) and 42 12-year-olds (19 girls and 23 boys). The former group have an average age of 116 months or 9;8 and the latter 150 months or 12;6.

Table 5.6: Russian subjects

	No of participants	**Mean Age**	**Range**
9-year-olds	32 (14F, 18M)	9;8	9;1 – 10;4
12-year-olds	42 (19F, 23M)	12;6	12;2 – 13;1

The children's L1 was always Russian as it was their home language. Russian was the mother tongue of all parents except for one father whose first language was Armenian. As seen in Table 5.7, children come from families of high socio-economic status as indicated by the parents' education. The vast majority of the parents (> 90%) have postgraduate education while only a few have professional training (7/147), and one reported elementary school while his or her partner was reported as postgraduate.

Table 5.7: Overview table of background variables extracted from the parental questionnaires of Russian children

	9-year-olds	**12-year-olds**
Cultural capital	Moscow	Moscow
Parents' L1	Russian	Russian
Home Language	Russian	Russian
Parental education	No degree: 4/64 (6%) Undergraduates: 59/64 (1%) Postgraduates: 59/64 (92%)	No degree: 4/83 (5%) Undergraduates: 5/83 (6%) Postgraduates: 74/83 (89%)

5.2.5 Length of Exposure & Input

Length of exposure in this study is operationalised as the number of years of instruction in EF which was one of the main criteria for 9-year-olds and 12-year-olds to participate in this study. However, the academic hours of attendance for each child in EF will also be considered as a second measure of input.

We saw in Section 5.2.1 above that for Russian children teaching hours increase with proficiency while in China the hours of classes remain constant for all levels. However, in Shanghai, EF offers a number of mini intensive courses over holidays which are not available in Moscow. Table 5.8 summarises the mean academic hours for Russian and Chinese age groups. As can be seen, there seems to be a considerable difference between the younger groups' mean academic hours of attendance. The older groups are more comparable.

Table 5.8: Mean academic hours and ranges of attendance in EF by both Chinese and Russian children

Academic hour= 40 mins	**CH_9y.o.**	**CH_12y.o.**	**RU_9y.o.**	**RU_12y.o.**
Academic hours in EF	880 [384 – 1455.5]	821 [546 – 1065]	592 [242 – 697.5]	804 [405 – 921.5]

Finally, we also need to consider English classes in day schools. In Shanghai, children start English at the age of 6 (primary 1) and receive around 4 hours of instruction per week on average (Tan, 2012) while the hours increase in secondary school to 5 or 6. In Moscow, children start learning English at day schools at 8 for 3 academic hours.

5.2.6 Proficiency

Turning to children's proficiency, a vocabulary task was administered, and the CEFR (Common European Framework for Reference for languages) levels of the classrooms each child attended at time of testing was requested from EF to get an even better-informed understanding of their proficiency level.

5.2.6.1 Word Finding Vocabulary Task

I administered the Word Finding Vocabulary Task (Renfrew, 1995) that aims to capture children's productive knowledge of L2 vocabulary. This is a standardised task and was originally developed for clinical purposes. The reason for employing this test is to get an indication of the children's vocabulary size as vocabulary size is often used to indicate differences in language development.

Children are asked to name the picture they see on a computer screen presented item by item on a PowerPoint presentation. There are in total 50 test items. In this study, all children were exposed to all items. At least 25 words out of the 50 exist in their books up to the A2 (CEFR) level.

To score, I took as correct answers the target answers provided by the test. Since it is a standardized task, I wanted to keep the scoring constant for all children so that comparisons would be fair among children. In some cases, children provided words close in meaning to the target ones such as 'joker' instead of 'clown', which were not counted as correct. A lenient scoring would not change the results much and thus I kept the strict scoring.

Table 5.9 summarises the mean vocabulary scores and ranges across groups. As can be seen, the group of Chinese 9-year-olds had a mean score of 17/50 (range: 9 – 27), while the group of Chinese 12-year-olds had 19,5/50 average score (range: 11 to 33).

As for the Russian children, the group of the 9-year-olds produced on average 18/50 words (range 9 – 27), with the Group of 12-year-olds scoring 24,5/50 (17 – 33).

Hence, the average scores of younger children in both countries are very similar, while there seems to exist a difference between the groups of older children.

Table 5.9: Group average scores (standard deviations) and their ranges in brackets concerning children's scores on the Renfrew Word Finding Vocabulary Task

	Chinese	**Russian**
9-year-olds	17/50 (4.55) [9 – 30]	18/50 (5.3) [5 – 27]
12-year-olds	19,5/50 (5.3) [11 – 33]	24,5/50 (4.8) [17 – 33]

To test whether the differences in means between the groups were significant, I carried out an Analysis of Variance (one-way independent-measures ANOVA)) using the R software (R Core Team, 2014) with age and L1 groups as the independent variable and Renfrew scores as the dependent variable.

I first checked whether ANOVA's assumptions were satisfied. The assumption of independent observations was met. I then checked for outliers visualising the data in boxplots; no outliers were found. The assumption of equality of variances was also met as Levene's test gave a non-significant result, $F(3, 143)=0.4906$, $p=0.6894$. Finally, with respect to the normality of the data assumption, the group sample sizes were all above 30 and according to the central limit theorem the sampling distribution tends to be normal if the sample is large enough (n>30). In fact, the Shapiro-Wilk normality test produced a p-value= 0.05506, that is a p-value>.05 (borderline) implying that the distributions of the data are not significantly different from normal distribution.

I then ran the main analysis using the aov[15] function. ANOVA results showed that there was a significant difference between the group means, $F(3, 143)= 17.43$, $p< 0.001$. I thus ran

[15] aov() function is an object performing one-way ANOVA in R.

a post-hoc test to see which group(s) differ. I used the Bonferroni method as my sample sizes were unequal but they had equal variances. The output revealed that younger children are not significantly different from each other (CH_9 – RU_9: p=1.00000). As for older children, Russians had significantly higher scores on the Renfrew task than Chinese counterparts (RU_12 – CH_12: p(0.00025)<.001. Russian 12-year-olds did significantly better than all the other groups, while Chinese 12-year-olds did not differ from the Chinese 9-year-olds nor from the Russian 9-year-olds on this task.

The Renfrew task then revealed a proficiency discrepancy between the older groups of the two L1 cohorts. Now, there are three possible reasons for this attested discrepancy in performance. First, learners' performance on vocabulary may have been affected by their L1; the linguistic proximity of Russian to English as opposed to Chinese. Second, the fact that Russians start learning English at their day schools later than Chinese when perhaps children are 'readier' for learning L2 vocabulary having better memory skills and learning strategies. Third, the education level of parents in Moscow which is higher than the education level of the parents in Shanghai; there are more postgraduates in Moscow while in Shanghai most parents hold a university degree. To test statically whether parental education differs among the two L1 groups having an effect on vocabulary, I ran a multiple linear regression analysis with dependent variable the vocabulary scores and as independent variables the L1 group and the parental education variables in interaction. I considered as parental education the highest education of any parent of the same household[16]. Results showed only a significant effect of L1 (p<0.001) with Russians doing significantly better than Chinese -which we already knew- but no significant interaction between the two, nor a main effect of parental education. This is probably due to the fact that the majority of the 12-year-olds' parents are educated holding at least a degree while there is no much variability to shed light on the potential effect of this variable[17]. Thus, this explanation cannot hold. I further discuss the learners' proficiency in the Discussion chapter.

Let us see now whether the proficiency discrepancy between 12-year-olds revealed in Renfrew is further confirmed by the CEFR level classes children attend.

---

[16] In the parental questionnaire, the question about the education level did not specify mother and father but 'your education' and 'your partner's education'. As such I could not consider maternal education or paternal education. I could also code the education levels of both parents and add the result in a new variable but the levels would be many and the interpretation would be difficult. What I did instead was to measure parental education considering the highest education of any parent as described in the text and code it using three levels: no university education: 0, university degree: 1, postgraduate degree: 2.

[17] Based on these results, I will not consider the parental education level as a variable in further analyses.

5.2.6.2 Class CEFR Level

Table 5.10 shows the proficiency levels considering the class children attended at the time of testing. As can be seen, there is variation in proficiency: there are 14 Russian 9-year-olds who attend an A1 CEFR level class while in the group of Chinese 9-year-olds there were only 5 with the majority (29) being in an A2 level class. It seems then that Russian 9-year-olds may have overall lower proficiency than their Chinese counterparts despite their Renfrew Vocabulary scores. On the other hand, older children from both L1 cohorts seem to attend classes of the same level which contrasts Renfrew test results presented above.

Table 5.10: Proficiency levels corresponding to the class children attended at time of testing within groups

	Chinese	**Russian**
9-year-olds	5 kids: A1 29 kids: A2 5 kids: B1	14 kids: A1 13 kids: A2 5 kids: B1
12-year-olds	31 kids: B1 3 kids: B2	39: B1 3 kids: B2

The picture we get from considering these proficiency measures is somewhat complicated; younger children although do not seem to differ with respect to the Renfrew task, the class CEFR level they attend seems to show proficiency differences. Then, older Russian children although attending the same CEFR level class with Chinese do better in the Renfrew task.

As Renfrew gave us different results than the class CEFR levels, I will later consider further measures of proficiency when analysing children’s narratives to clarify further the question of the proficiency of each group.

5.2.7 Research materials

The same tasks were used in both locations, Shanghai and Moscow. All the tasks were production tasks which would give us information on the kinds of errors with verbal morphology L2 children make and what stages they go through in their acquisition process. Furthermore, production provides a direct way to find out whether a particular linguistic phenomenon be it a morpheme or a structure has been acquired (Ionin & Zyzik, 2014).

However, in free production a child could perhaps avoid certain structures, which is why elicited production may be particularly enlightening.

Past tense and present 3SG-agreement morphemes were assessed through the probes of an elicited production task, the Test of Early Grammatical Impairment (Rice & Wexler, 2001) as well as through narratives.

The TEGI was used in various previous studies so this would allow us make comparisons with children or adults from other studies. Then, in free production, instances of use of an actual morpheme may be very few and the obligatory contexts difficult to determine. In addition, children may use memorized formulas and/or only frequent morphemes. The elicited production tasks of TEGI specify the context as either present/habitual or past allowing to better interpret children's answers.

However, I also used free spontaneous production to obtain a more complete picture of the children's knowledge of verb morphology. The elicited production task such as TEGI has its own limitations; for example, children especially older ones may realize what the test asks for and apply a rule consistently and consciously. Therefore, both modalities are necessary to have a clear picture of children's production of morphological markers.

All the tasks were child-friendly and child appropriate/suitable and were already piloted as well as employed in previous studies confirming the validity of their content.

I present each task in the next section.

5.2.7.1 Test of Early Grammatical Impairment (TEGI)

The TEGI (Rice & Wexler, 2001) is a battery of elicited production tasks which includes production probes for third person singular (3SG) and past tense. The TEGI capturing 3SG-agreement consisted of 10 test items following a practice item. The child saw a picture on a laptop screen depicting a person (e.g. a teacher, a dentist, a dad) and the experimenter asked the child what each person does. For this set of items, the targeted form in the answer needs the verb inflected with 3SG -*s*. See an example in Figure 5.1 below:

Figure 5.1: Example of a test item of the task (TEGI 3SG). In this task, the experimenter would say: 'Here is a teacher. Tell me what a teacher does.' Target answer: A teacher teaches

I administered the test following the TEGI guidelines with one exception; contrary to what the test suggested, I did not provide the correct answer in the example item if the child gave a wrong answer. The reason for this is the possibility that it could favour older children who are better test-takers and may have more practice than younger children with these types of tasks. Note however, I did correct if s/he described the picture using the present continuous. For example, if the child said: '*A teacher is teaching.*', the experimenter would say '*Don't tell me what this teacher (pointing to the picture) is doing. Tell me what a teacher does. Any teacher.*', or '*Don't describe the picture. Tell me what any teacher does.*'

To elicit the past tense, I used the TEGI probe for past tense. The task included 2 trial items and 18 test items; 10 targeting regular verbs and 8 targeting irregular verbs. Children were shown two pictures on a computer screen. The experimenter would describe the first picture and ask the child to describe the second one. Figure 5.2 displays an example test item.

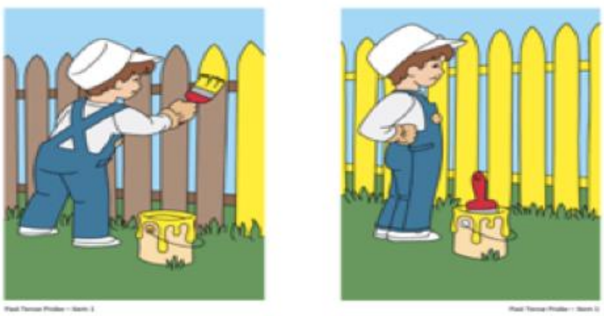


Figure 5.2: Example of a test item of the TEGI Past task. In this task, the experimenter would say: 'Here the boy is painting the fence (pointing at the first picture). Now he is done (pointing at the second picture). Tell me what he did. He …' Target answer: He painted (the fence)

As with the previous test, the procedure indicated by the TEGI manual was followed with one exception; the correct answer in the example item was not given to children.

5.2.7.2 Multilingual Assessment Instrument for Narratives (MAIN)

I also targeted tense production in narratives. For this purpose, the Multilingual Assessment Instrument for Narratives (henceforth MAIN) by Gagarina et al. (2012) were used. The child was given a sequence of pictures depicting a story and asked to say the story. In Figure 5.3, an example story with the sequence of pictures of the MAIN battery is illustrated.

Figure 5.3: An example story of the narratives: 'The dog story'

Each story included six pictures all having parallel structures. In the first picture the main protagonist was depicted and the second protagonist always appears in the second picture. Two of the stories (i.e. Cat, Dog) included a main protagonist (e.g. a cat or a dog) while the other two (i.e. Baby Birds, Baby Goats) had two main protagonists (two baby birds or two baby goats), thus controlling for plurality across the stories. Each six-picture sequence was designed to illustrate three episodes each of which included an initiating event, a goal of the protagonist, an attempt to achieve the goal, an outcome of the attempt and the reaction to the outcome. So, each story should start with the setting providing the time and place reference followed by three episodes. Each episode is developing in parallel to another one, thus concurrent events are illustrated.

Two modes were employed in this study; the retelling mode, when the child would be asked to tell the story (either the Cat story or the Dog story) after listening to it (see Appendix B for recorded stories), and the telling mode in which the child was asked to tell a story using their own words (Baby Birds or Baby Goats). In the retelling mode, the child would listen to the story as was recorded by a native English speaker. Administration of the task followed the MAIN's guidelines.

5.2.7.3 Parental questionnaires

Parental questionnaires were administered to gather information about children's biodata as well as their families (Appendix A.4). The parental questionnaires were translated into Chinese and Russian in order to make them easy for parents to understand.

The main aim of the questionnaire was to gather information on language input and educational background. The parental questionnaire consisted of 8 Sections. The first asked information about the child's basic biodata, that is name, date of birth and gender. The second section asked questions about the age of onset. Questions included when the child started learning English in EF, when in their day school, if s/he had exposure to the language before joining EF, and if yes, what type of learning experience that would be (e.g. private school, private tutor, bilingual kindergarten). The third section asked information about the child's exposure to and use of English (e.g. time devoted to homework, multimedia use, travelling). The fourth section concerned the languages used at home. Section 5 sought information about the parents' attitudes towards their child's English language learning as well as the child's motivation and enjoyment of the classes (e.g. How important is it for you that your child learns English? Does your child seem to enjoy learning English?). I also included a couple of questions in Section 6 on whether the children had any language problems to ensure all children were typically developing. Some questions concerning the development of literacy were also added in Section 7. Specifically, questions were: 'At what age did your child learn to write and read in Chinese/Russian?' and 'Does he/she know how to read and write in a language other than Chinese and English?'. Finally, I asked about the parents' educational background in order to determine the socioeconomic situation of the family. The questionnaire was two-pages long.

Parental questionnaires were translated back into English by a native Chinese and a native Russian speaker and the data were entered into excel files. Names were deleted and codes were given to each child.

5.2.8 Data collection procedure

The very first step prior to the data collections was to apply for ethical approval for this project, which was given by the Ethics Committee for the School of the Humanities and Social Sciences of the University of Cambridge (See Appendix A.1).

Subsequently, and in order to ensure the feasibility and the appropriateness of the tests, I conducted a pilot study. The pilot study took place in Shanghai in March 2017 and during two weekends 43 children participated in the research. After refining the tasks and the criteria of the design, the main data collections in Shanghai and Moscow took place; in November-December 2017, and in October-November 2018, respectively.

Prior to the data collections, consent forms (Appendix A.2.1- A.2.2), parent information letters (Appendix A.3) as well as parental questionnaires translated in either Chinese or Russian were distributed to parents. Consent forms and parental questionnaires were then collected when the experimenter visited the school. This process was facilitated by EF school directors.

Children were tested in their EF schools by myself and the whole process was recorded for later transcription. Interviews took place in EF schools in Shanghai and Moscow, and children were seen on an individual basis and in one session. We tried to ensure that classrooms would be quiet and we managed it to a large extent, however, in certain cases this was not possible. In any case, the overall quality of the recording is high.

The testing process was the same in both data collections. The experimenter would sit opposite, or next to the child depending on the room's layout to help better interaction with the child and allow for easy eye contact. There were additional reasons for this seating arrangement which had to do with the tasks. For the narratives, for instance, the experimenter was not supposed to see the pictures -especially in the telling mode- as the aim was to let the child believe that only s/he could see the pictures. This way effects of shared knowledge would be avoided. The whole testing session lasted on average 45 minutes. All the children, and especially the younger ones, were encouraged by the experimenter throughout the testing sessions and regardless of their performance.

At the start of the session, the experimenter would get to know the child's name and age and ask questions such as 'How are you?', 'How was your day today?', 'Did you have a class before?', etc. so as to establish rapport with the child and make him/her feel comfortable. The experimenter would then introduce herself giving the child her name and then explaining that they would do some tasks together. She also tried to make the children feel comfortable by telling them that this would not be an exam session and that there would be no marks or grades to worry about.

The testing process would then start following this order of tasks (approximate average time needed per task is given in parentheses):

1. Narratives 1 – retelling mode (5')
2. Wh-questions (5')

3. Narratives 2 – telling mode (5')

4. Relative clauses (10')

5. TEGI 3SG (5')

6. TEGI Past (5')

7. Renfrew task (5')

Note that the tasks for the wh-questions and the relative clauses will not be discussed in this work due to space limitations. However, as they were part of the testing battery need to be mentioned.

The first task would be the MAIN (Narratives). The retelling narrative MAIN was first to ensure that the child will not be too tired to tell a story and might be an easier task for the child to start the session with. This was followed by the elicitation of wh-questions. The second part of the MAIN (Narratives) was third, in which the child was asked to tell a story without listening to it. After the narratives, the elicited production task for the relative clauses followed. The next task is TEGI; first the probes for 3SG-agreement, then the probes for past tense. Finally, the Renfrew task was administered. The Renfrew was much liked by the children and perceived as a more relaxing task towards the end of the session.

At the end of the testing session, which was recorded, the experimenter would thank the child and would give him/her a certificate of participation. In addition, they would all be encouraged with phrases such as "You did a very good job.", "Well done", etc, regardless of their performance.

The whole testing session was in English although instructions were available in Chinese/Russian if participants could not understand. The vast majority of participants did not use the translations.

5.2.9 Data transcription

To audio-record the data I used a professional recording device to ensure the quality of the sound files. After each testing session, I transferred the sound files to the laptop, renamed them by codes and organized them according to the group each child was placed in. I transcribed the data orthographically.

5.3 Summary

In this chapter, I presented the current study and specifically discussed the research questions and hypotheses of this work as well as the methodology used to address them. In a nutshell, this thesis seeks to answer questions regarding how children learn particular linguistic properties at certain ages focusing on the impact of age (of onset) and the influence of their L1 over the L2. All phenomena were assessed through production tasks which were audio-recorded and transcribed.

In the next chapter, I present results of the TEGI probes for 3SG-agreement and Past tense and in Chapter 7, I discuss results on the same markers in narratives.

CHAPTER 6

RESULTS ON FINITENESS MARKING IN ELICITED PRODUCTION TASKS

6.0 Introduction

In the previous chapter, I presented the research hypotheses and the relevant predictions as well as the methodology of this study.

Focusing on the acquisition of finiteness, I asked:

- whether L2 children with age of onset after 4 will pattern with L2 adults in the domain of inflectional morphology
- whether there is a gradual decline in performance as a function of older age of onset
- whether Russians whose language has the relevant features will outperform Chinese who lack such features, and
- whether the interpretability status of the features are determinants of their difficulty in acquisition.

In this chapter, I present the results of the TEGI elicited production tasks. I discuss each test and each L1 cohort separately and provide both group results as well as individual results. Before doing so, the coding and scoring of the tests will be presented.

6.1 Data coding and scoring of the probes

In this section, I present the coding method used for children's answers in TEGI. With my aim being to investigate children's performance including error patterns, I did not employ TEGI's coding as it would not be enlightening of children's performance leaving a large number of non-target answers unscorable. (See Appendix C for tables using the TEGI's scoring guidelines.)

6.1.1 Coding and scoring of the TEGI 3SG

I first consider TEGI for 3SG-agreement. The scoring method was as follows; correct answers were considered those which had a 3rd person singular subject and a verb marked with 3SG -*s* (e.g. '*A father/He plays with his children.*'). In case the child used a 3rd person singular subject and did not mark the verb with 3SG -*s*, the answer was scored as incorrect. I further calculated instances of periphrastic marking which included instances of (grammatical and ungrammatical) *be + verb-ing, be + verb, modal verb + verb, does + verb*. Examples of this category included instances such as: **He is play, she is playing, she can play; he does play*. Any other response with use of another tense (e.g. simple past), subject in plural, progressive participle -*ing*, *does* + *verb-ing* or noun, and copula '*be'* were classified as 'other'. Finally, 'no responses' were also recorded. Table 6.1 presents the scoring method along with examples.

Table 6.1: Scoring schema used for TEGI 3SG along with examples

Scoring	**Structure**	**Example**
Correct	(Subj in 3SG +) verb-s	*He plays*
Incorrect	(Subj in 3SG +) verb-Ø	*She dance*
Periphrastic marking	is verb-ing	*He is fixing*
	is verb-Ø	*He is play*
	modal verb + verb	*He can paint*
	does + verb	*He does play*
Other	does + verb-ing	*He does painting*
	verb-ing	*She flying*
	past tense	*He saved*
	copula be	*He is strong*
	to-verb	*He to play*
	other types of errors	*He does to playing*
No response	no verb or no answer at all	*He ….*

To establish the proportion of correct answers, I summed up all the correct instances of 3SG marking and divided that number by the total number of test items. The same process was followed in order to measure the proportion of all other categories.

6.1.2 Coding and scoring of the TEGI PAST

I followed a similar method for TEGI Past. Correct answers were those inflected with *–ed* (e.g. *She cleaned*) or the correct suppliance of the irregular form (e.g. *He made*). Use of a different regular verb other than the target regular, but correctly inflected with *–ed* was also scored as correct (e.g. *tidied* instead of target *cleaned*). The same holds for irregular verbs (e.g. *built* instead of *made*). Instances of overregularization, that is, answers with irregular verbs inflected as regular and marked with *–ed* were scored separately (but would then be added up with the correct answers). Omission of inflection in this past context was taken as incorrect (e.g. *She clean-∅* or *He make*). Periphrastic marking was also recorded here including again (grammatical or ungrammatical) '*be + verb-ing, be + verb, be + inflected verb, modal verb + verb, did + verb'*. Examples of this category were instances such as '**The girl is/was clean her room, the boy is jumping, the boy can jump, the girl did clean.*' Use of another tense (such as present simple, present perfect, past perfect), progressive participle as the main verb, '*did'* + *verb-ing* or noun, '*be'* as a main verb, use of verbs such as *hit, put, run* for which we cannot be sure whether they were marked for past tense, and use of another type of verb (e.g. regular instead of irregular) were classified as 'other'. There were very few cases that could not be classified in a certain category such as '… *to brushing* ...', '… *did is paint ...',* which were marked as 'Other'. Finally, 'no responses' were also recorded. Table 6.2 illustrates schematically the scoring used for children's answers in TEGI past and relevant examples.

Table 6.2: Scoring schema used for TEGI Past along with examples

Scoring	**Structure**	**Example**
Correct (regular)	verb-ed	*She planted*
Incorrect (regular)	verb-Ø	*She plant*
Correct (irregular)	verb in past form	*She dug*
Correct (overregularization)	verb-ed	*She digged*
Incorrect (irregular)	verb	*She dig*
Periphrastic marking	be verb-ing	*He is giving/He was playing*
	be verb-Ø	*He is play /He was play*
	be + inflected verb	*He is played*
	modal verb + verb	*He can play*
	did + verb	*He did play*
Other	did + verb-ing	*He did painting*
	verb-ing	*He painting*
	present simple	*He paints*
	copula be	*He is …*
	verbs like hit, put, run	*She hit the ball.*
	use of other type of verb	*She left the box (instead of lifted)*
No response	no verb or no answer at all	*He … .*

To calculate the proportion of correct answers, I added up all the correct instances of past tense marking (i.e., regular + irregular + overregularization) and divided that sum by the total number of test items. In the same way I calculated proportions of past incorrect, periphrastic marking, 'other', and 'no response' categories. The denominator was always the total number of test items except for the cases when I wanted to see only the regular or the irregular contexts. In order to calculate the correct inflection of regular verbs only, I divided the number of instances of correctly inflected regular verbs (with *-ed*) by the number of contexts where a regular verb was required. I calculated the percentage of correctly supplied irregular verbs in the same way. To calculate the percent of overregularization, I took the sum of the instances of overregularization and divided it by the total number of irregular verbs, that is, the contexts where an irregular verb would be required.

6.2 TEGI 3SG probe verbs

TEGI 3SG probes consisted of 10 test items providing habitual present tense contexts. In TEGI 3SG, probes do not require any specific verbs and are open to any kind of answer. However, considering the drawings which show people of various professions or doing certain activities, we are likely to find the following set of verbs:

1. A dentist fixes/makes our teeth.
2. A police officer protects the city.
3. A firefighter puts out the fire.
4. A pilot flies a plane.
5. A painter paints buildings.
6. A baseball player plays baseball.
7. A nurse helps people/patients.
8. An astronaut goes to the moon.
9. A dad plays with his children.
10. A dancer dances.

Again, children can use any verb they want as their choices are not restricted. I start by providing results for the Chinese children and then turn to the Russian children. Results will be presented in terms of raw scores as well as percentages.[18]

6.3 Results on TEGI 3SG

6.3.1 Results on TEGI 3SG for Chinese children

Thirty-seven of the 39 Chinese 9-year-olds included in the research participated in this task. Two children were excluded as they failed to do the task because of lack of time or lack of understanding of the instructions (likely a result of low proficiency). All 34 participants of the 12-year-olds did the task. Thus, there were 37 x 10 obligatory present tense habitual contexts for the group of younger children and 34 x 10 for the group of older children.

Table 6.3 shows the results for the two groups. Percentages are rounded up to the closest integer.

As can be seen, the group of younger children correctly mark the verb with 3SG -*s* 10% of times; the older children mark to 42% accuracy. Both groups produce bare forms, to the same extent; younger children: 40% and older children: 42%. The younger children use

[18] Statistical analyses are presented in Chapter 9 where I consider results all together and explain the linear mixed effects models used for all analyses.

periphrastic marking at 37% which older children do not use as much (9%). Finally, both groups produce ‘other’ responses to similar extents (11% and 7% for younger and older children respectively).

What these results show is that the bare forms produced by older children (42%) did not reduce (younger children: 40%) whereas the correct forms increased from the 9-year-olds to the 12-year-olds. The increase in accuracy then is due to a decrease in periphrastic marking produced by the younger children.

Table 6.3: Chinese children’s performance on TEGI 3SG in raw scores and percentages

	Correct 3SG *-s*	**Incorrect 3SG Ø**	**Periphrastic marking**	**Other**	**No response**
9y.o. (n= 37)	38/370 **10%**	149/370 **40%**	138/370 **37%**	42/370 11%	3/370 1%
12y.o. (n= 34)	142/340 **42%**	142/340 **42%**	31/340 **9%**	25/340 7%	0/340 0%

Given their surprisingly large quantity, the next step was to try to find patterns within the periphrastic marking category. As we have already discussed above, four categories were created: ‘*be* + *verb-ing*’, ‘*be* + *verb*’, ‘*modal* + *verb*’ and ‘*does* + *verb*’. The most common pattern found was the ‘*be* + *verb(-ing)*’ construction which accounted for 27% of the younger children’s production and for 9% of the older children’s production. Older children only used this structure (‘*is* + *verb(-ing)’)* and did not use modals or *‘does* + *verb’* at all. An example of this structure is ‘*The dentist is fix teeth.*’ instead of ‘*A dentist fixes teeth.’* or the grammatical but not target: ‘*The dentist is fixing teeth.*’. Other instances of periphrastic marking included instances of ‘*modal* + *verb’* and ‘*does* + *verb’*, however we will not discuss these further as they amount to very few cases compared to the ‘*be* + *verb(-ing)*’ pattern. In Table 6.4, below, both the raw numbers and the percentages of the use of the ‘*be* +*verb(-ing)*’ structure are given. Note that the ungrammatical use of this structure, that is, the use of ‘*be* + *bare verb’* is much more frequent than the grammatical but not target structure especially in younger children’s production.

Table 6.4: Children's use of '*be + verb*' (including both grammatical '*be + verb-ing'*, and ungrammatical '*be + verb'*) in the TEGI 3SG

	Grammatical 'be + V-ing' + Ungrammatical 'be + stem verb'	**Total test items**
9y.o.	(24 + 77)/370 **27%**	370 (100%)
12y.o.	(15 + 16)/340 **9%**	340 (100%)

To ensure that these '*be + verb(-ing)*' forms did not appear with certain verb types only, an item analysis was carried out. As can be seen below, in Table 6.5, this does not seem to be the case as there are not large deviations in the instances of these forms across verbs/test items.

Table 6.5: Item effects analysis for 'be + verb (-ing)' forms in TEGI 3SG

be forms	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
9y.o.	13	7	9	12	11	10	8	5	13	13
12y.o.	2	3	3	3	3	3	3	3	4	4

I also found some instances of the present participle without the auxiliary *be* (i.e. *verb-ing*). This constituted the main type of response included in the 'other' category in both groups (22/370 for the 9-year-olds, 20/340 for the 12-year-olds); for example, '*A baseball player playing baseball*.' Other types of responses categorised as 'other' would be copula '*be'* as in '*He is strong'*, use of the simple past (very rare) as in '*He helped for teeth*' '*does + verb-ing'* as in 'She *does painting'*.

Let us consider individual results. Table 6.6 shows individual results for 3SG. Results are sorted by the highest score of correct answers to the lowest.

Table 6.6: Individual results for Chinese 9-year-olds on TEGI 3SG

	ID	Correct 3SG	Incorrect (bare) form	Periphrastic marking	Other	No response	Total test items
1	CH_GA_05	10	0	0	0	0	10
2	CH_GA_04	6	3	1	0	0	10
3	CH_GA_32	5	3	0	0	2	10

4	CH_GA_06	4	5	0	1	0	10
5	CH_GA_09	4	3	0	3	0	10
6	CH_GA_22	4	5	1	0	0	10
7	CH_GA_11	2	5	3	0	0	10
8	CH_GA_12	1	1	6	1	1	10
9	CH_GA_31	1	6	0	3	0	10
10	CH_GA_34	1	4	1	4	0	10
11	CH_GA_01	0	2	6	2	0	10
12	CH_GA_02	0	1	8	1	0	10
13	CH_GA_03	0	7	3	0	0	10
14	CH_GA_07	0	2	8	0	0	10
15	CH_GA_08	0	0	9	1	0	10
16	CH_GA_10	0	9	1	0	0	10
17	CH_GA_13	0	2	2	6	0	10
18	CH_GA_14	0	7	1	2	0	10
19	CH_GA_15	0	0	9	1	0	10
20	CH_GA_17	0	0	9	1	0	10
21	CH_GA_18	0	5	4	1	0	10
22	CH_GA_19	0	9	0	1	0	10
23	CH_GA_20	0	1	7	2	0	10
24	CH_GA_21	0	2	7	1	0	10
25	CH_GA_23	0	2	8	0	0	10
26	CH_GA_24	0	10	0	0	0	10
27	CH_GA_25	0	1	8	1	0	10
28	CH_GA_26	0	0	10	0	0	10
29	CH_GA_27	0	9	1	0	0	10
30	CH_GA_28	0	9	1	0	0	10
31	CH_GA_30	0	10	0	0	0	10
32	CH_GA_33	0	0	10	0	0	10
33	CH_GA_35	0	2	4	4	0	10
34	CH_GA_36	0	9	0	1	0	10
35	CH_GA_37	0	7	0	3	0	10
36	CH_GA_38	0	6	3	1	0	10
37	CH_GA_39	0	2	7	1	0	10
	TOTAL	**38**	**149**	**138**	**42**	**3**	**370**

Table 6.6 demonstrates that only 10 out of the 37 children have produced any 3SG -*s*. There were also only 10 children who did not produce any form of '*be* + *verb*(-ing)', that is, 27 children did produce at least 1 form of this structure. In addition, 4 children produced only

periphrastic marking. This picture clearly shows that this is not a random performance but something we need to focus on and explain. Bare forms were more dispersed among children.

Turning to the group of 12-year-olds, their performance in terms of individual results on TEGI 3SG is illustrated in Table 6.7 (data sorted from largest score of correct production of 3SG to the lowest). Note that the column 'No response' was omitted here because there were no 'no responses'.

Table 6.7: Individual results for Chinese 12-year-olds on TEGI 3SG

	ID	Correct 3SG	Incorrect (bare) form	Periphrastic marking	Other	Total test items
1	CH_GB_16	10	0	0	0	10
2	CH_GB_09	9	1	0	0	10
3	CH_GB_14	9	1	0	0	10
4	CH_GB_21	9	1	0	0	10
5	CH_GB_08	7	3	0	0	10
6	CH_GB_17	7	2	0	1	10
7	CH_GB_20	7	3	0	0	10
8	CH_GB_30	7	3	0	0	10
9	CH_GB_34	7	3	0	0	10
10	CH_GB_01	6	4	0	0	10
11	CH_GB_05	6	3	0	1	10
12	CH_GB_06	6	4	0	0	10
13	CH_GB_12	6	4	0	0	10
14	CH_GB_15	6	4	0	0	10
15	CH_GB_13	5	4	1	0	10
16	CH_GB_28	5	4	0	1	10
17	CH_GB_32	5	5	0	0	10
18	CH_GB_02	4	6	0	0	10
19	CH_GB_24	3	7	0	0	10
20	CH_GB_27	3	7	0	0	10
21	CH_GB_07	2	1	3	4	10
22	CH_GB_23	2	7	0	1	10
23	CH_GB_26	2	7	0	1	10
24	CH_GB_29	2	8	0	0	10
25	CH_GB_33	2	7	0	1	10
26	CH_GB_18	1	5	0	4	10
27	CH_GB_19	1	9	0	0	10
28	CH_GB_22	1	2	7	0	10

29	CH_GB_25	1	8	1	0	10
30	CH_GB_31	1	9	0	0	10
31	CH_GB_03	0	0	10	0	10
32	CH_GB_04	0	2	0	8	10
33	CH_GB_10	0	1	8	1	10
34	CH_GB_11	0	7	1	2	10
	TOTAL	**142**	**142**	**31**	**25**	**340**

It is again only 1 child that gave 10 correct responses out of the 10 in this task. Excepting also three children who scored 9/10, the rest scored between [0-7] still showing great fluctuation, even though the majority of the children now got at least one correct answer. As for the periphrastic marking, this was produced only by 7 children which shows that this strategy has mostly been abandoned by older children.

What we have seen so far is that the Chinese children have problems with the tense and agreement marker -*s* as they produce it to a very limited extent at these early stages of learning despite the structure having been taught. Individual results further confirm what the group percentages showed us. Older children's performance is less diverse across children, while younger children's performance is not distributed merely between correct and incorrect answers but they also do something different; they employ a strategy with the most common pattern being '*be + verb(-ing)'*.

Let us turn to the Russian children in order to see whether their performance is similar or not.

6.3.2 Results on TEGI 3SG for Russian children

All 32 9-year-olds participated in the task but one child was excluded as some of his/her answers were unintelligible and some were missing. Thus, a total of 31 9-year-olds were included in the research as well as all 42 older children. In Table 6.8, the results are given.

As can be seen, the 9-year-olds produced 3SG -*s* to an extremely limited extent (2%), that is, 3SG-agrement morphology is almost non-existent. 12-year-olds did better marking the verbs with 3SG -*s* at 27% of their responses. There is an increase then in correct answers from 9 to 12 years of age while there is a decrease of bare forms from 30% to 19,5% for older children. Both younger and older groups -the younger children slightly more- use the periphrastic marking to a great extent which accounts for 42% and 38.5% of their total answers respectively. Finally, the 'other' category is also quite high for both groups (9y.o.: 23%; 12y.o.:15%).

Table 6.8: Russian children's performance on TEGI 3SG in raw scores and percentages

	Correct 3SG	**Incorrect 3SG**	**Periphrastic marking**	**Other**	**No response**
9y.o. (n=31)	7/310 **2%**	94/310 **30%**	131/310 **42%**	71/310 23%	7/310 2%
12y.o. (n=42)	114/420 **27%**	83/420 **19,5%**	162/420 **38,5%**	61/420 15%	0/420 0%

As before, the periphrastic marking category was put under the microscope to reveal patterns in Russian children's performance. Following a similar procedure as above, I found again that '*be + verb(-ing)*' was the main pattern. In Table 6.9, the exact number of instances of the '*be + verb(-ing)*' can be found for both groups. The 9-year-olds produced both '*be + bare verb*' and '*be + verb-ing*' while 12-year-olds produced more '*be + verb-ing'* forms.

Table 6.9: Children's use of '*be + verb*' (including both grammatical '*be + verb-ing*', and ungrammatical '*be + verb'*) in the TEGI 3SG

	Grammatical 'be + V-ing' + Ungrammatical 'be + bare verb'	**Total test items**
9y.o.	(59 + 65)/310 **40%**	310 (100%)
12.y.o.	(91 + 36)/420 **30%**	420 (100%)

Again, data and particularly '*be + verb (-ing)*' forms were analyzed for item effects that could have triggered them, but as can be seen in Table 6.10, this is not the case, as these instances were not produced with only certain verbs but with all verbs to very similar extents.

Table 6.10: Item effects analysis for 'be + verb (-ing)' forms in TEGI 3SG

be forms	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
9y.o.	9	11	13	15	13	13	12	15	10	13
12y.o.	15	13	11	15	15	10	13	10	11	14

As for the children's other types of responses, again the main type would be the bare present participle (i.e. no aux.) for both groups. Use of copula '*be'*, '*does + verb-ing*', and past tense were rare types of responses.

Let us finally look in more detail at individual results to see whether the mean percentages accurately present the individual data. In Table 6.11, the individual results on TEGI 3SG for the 9-year-olds are presented (data sorted from largest score of correct answers to the lowest).

Table 6.11: Individual results on TEGI 3SG for Russian 9-year-old children

	ID	Correct 3SG	Incorrect (bare) form	Periphrastic marking	Other	Total test items
1	RU_GA_19	2	3	5	0	10
2	RU_GA_30	2	3	5	0	10
3	RU_GA_10	1	7	0	2	10
4	RU_GA_15	1	7	0	1	10
5	RU_GA_18	1	0	8	1	10
6	RU_GA_01	0	0	6	4	10
7	RU_GA_02	0	0	8	2	10
8	RU_GA_03	0	4	3	3	10
9	RU_GA_04	0	5	1	4	10
10	RU_GA_05	0	0	10	0	10
11	RU_GA_06	0	5	0	5	10
12	RU_GA_07	0	5	0	3	10
13	RU_GA_08	0	0	8	1	10
14	RU_GA_09	0	6	2	1	10
15	RU_GA_12	0	6	1	3	10
16	RU_GA_13	0	1	0	9	10
17	RU_GA_14	0	0	9	0	10
18	RU_GA_16	0	1	8	1	10
19	RU_GA_17	0	5	2	3	10
20	RU_GA_20	0	4	6	0	10
21	RU_GA_21	0	1	7	2	10
22	RU_GA_22	0	4	0	6	10
23	RU_GA_23	0	10	0	0	10
24	RU_GA_24	0	1	7	1	10
25	RU_GA_25	0	2	6	2	10
26	RU_GA_26	0	0	8	2	10
27	RU_GA_27	0	4	1	5	10
28	RU_GA_28	0	6	3	1	10

29	RU_GA_29	0	2	0	8	10
30	RU_GA_31	0	1	8	1	10
31	RU_GA_32	0	1	9	0	10
	TOTAL	**7**	**94**	**131**	**71**	**310**

It is evident that were only 5 9-year-olds who correctly marked the verb with 3SG -s but only 1-2 times out of the total 10 test items provided by the test. However, there were 23 children out the 31 participants who produced some form of ‘*be* + *verb(-ing)*’. Finally, bare forms were produced by most of the children to some extent.

Turning to individual results for 12-year-olds on the same test, these are illustrated in Table 6.12 (data sorted from largest score of correct answers to the lowest).

Table 6.12: Individual results on TEGI 3SG for Russian 12-year-old children

	ID	Correct 3SG	Incorrect (bare) form	Periphrastic marking	Other	Total test items
1	RU_GB_04	10	0	0	0	10
2	RU_GB_06	10	0	0	0	10
3	RU_GB_03	9	1	0	0	10
4	RU_GB_08	9	0	1	0	10
5	RU_GB_28	9	1	0	0	10
6	RU_GB_07	8	0	2	0	10
7	RU_GB_18	8	1	1	0	10
8	RU_GB_35	8	1	1	0	10
9	RU_GB_14	7	3	0	0	10
10	RU_GB_15	7	1	2	0	10
11	RU_GB_32	6	0	3	1	10
12	RU_GB_26	5	1	4	0	10
13	RU_GB_16	3	2	3	2	10
14	RU_GB_01	2	4	0	4	10
15	RU_GB_05	2	5	3	0	10
16	RU_GB_25	2	5	1	2	10
17	RU_GB_36	2	4	4	0	10
18	RU_GB_11	1	1	7	1	10
19	RU_GB_20	1	0	9	0	10
20	RU_GB_22	1	5	0	4	10
21	RU_GB_23	1	4	1	4	10
22	RU_GB_30	1	0	9	0	10
23	RU_GB_31	1	1	7	1	10
24	RU_GB_42	1	9	0	0	10

25	RU_GB_02	0	10	0	0	10
26	RU_GB_09	0	0	10	0	10
27	RU_GB_10	0	1	9	0	10
28	RU_GB_12	0	0	10	0	10
29	RU_GB_13	0	5	5	0	10
30	RU_GB_17	0	1	9	0	10
31	RU_GB_19	0	1	7	2	10
32	RU_GB_21	0	2	6	2	10
33	RU_GB_24	0	0	9	1	10
34	RU_GB_27	0	0	6	4	10
35	RU_GB_29	0	3	0	7	10
36	RU_GB_33	0	0	10	0	10
37	RU_GB_34	0	3	2	5	10
38	RU_GB_37	0	5	1	4	10
39	RU_GB_38	0	3	5	2	10
40	RU_GB_39	0	0	2	8	10
41	RU_GB_40	0	0	9	1	10
42	RU_GB_41	0	0	4	6	10
	TOTAL	**114**	**83**	**162**	**61**	**420**

Considering the individual results of the group of the older children, apart from 8 children who gave correct answers at least 8/10, 30 children gave less than 5/10 correct responses. Also, 32 of the older children used periphrastic marking and mainly '*be + verb-ing*'. Importantly, for those children who score above 7/10, instances of '*be + verb(-ing)*' is very rare showing that these children have mostly dropped this strategy. For many children though, it seems that this might be a later stage in their development as they still rely on this construction as the younger ones do.

To sum up, both Russian child groups have problems with inflectional morphology concerning 3SG-agreement, but especially for the group of younger children the bound morphology as in 3SG marking does not seem to have emerged at all. Children also do not produce only bare forms but mainly periphrastic marking which is a strategy used by both groups. '*Be + verb(-ing)*' is again the main pattern. However, a considerable number of the older children (8/42) seem to have overcome that stage and have mastered the use of the 3SG morpheme -*s*.

6.3.3 Summary

Discussing results from the TEGI 3SG, we saw that the groups of older children -both Chinese and Russian- outperformed the groups of younger children, being more accurate in marking the verb with 3SG *-s*. However, even the groups of older children have not yet mastered this property as their percentages of accuracy do not even exceed the 50% correct level. Individual results further confirmed the fact that 9-year-olds in both L1 cohorts have not learnt the 3SG marking which is only attested in very few children. Older children show variability in their responses extending to the whole range from 0 to 10 with again only few children performing at ceiling.

Children have attended English classes for five years while they have also been exposed to English in day schools; it seems then that this particular affix is quite difficult for L2 learners to acquire despite the fact that the relevant feature exists in one of their L1s.

Russians produce it even less than Chinese. This may seem unexpected considering the fact that Russian is a highly inflecting language which does mark verbs for Present tense, number, and person while Chinese being an isolating language does not mark verbs for tense/agreement. Thus, the reverse pattern was expected to emerge. I will come back to these findings later on in the discussion section in Chapter 10.

Third, both Chinese and Russian children -and especially the younger ones- use periphrastic marking and more concretely '*be + verb(-ing)*'. This structure is robust across L1 cohorts. However, while Chinese 12-year-olds seem to have abandoned this strategy, the Russian counterparts do not drop periphrastic marking.

Finally, we saw that both populations show similar patterns and types of 'errors', so apart from correct and bare forms as well as forms of periphrastic marking, they use the progressive participle, '*does + verb-ing'*, while the use of the simple past is rare in this context.

6.4 TEGI Past tense probe verbs

Before we see results on TEGI past tense, let us briefly consider the probe verbs. The task includes 18 probes; 10 eliciting past regular and 8 eliciting past irregular. Target verbs are shown in Table 6.13. In terms of lexical aspectual class, verbs are classified as accomplishments and achievements following Vendler (1967) and Shirai & Andersen (1995).

Table 6.13: Target verbs in TEGI Past

No of items	Target verbs	Regular/Irregular	Lexical class
1	painted (a fence)	regular	accomplishment
2	*caught (a ball)*	irregular	achievement
3	*made (a bird house)*	irregular	accomplishment
4	brushed (his hair)	regular	accomplishment
5	kicked (a ball)	regular	achievement
6	cleaned (her room)	regular	accomplishment
7	*wrote (a word)*	irregular	accomplishment
8	climbed (the ladder)	regular	accomplishment
9	jumped (-)	regular	achievement
10	*rode (the horse)*	irregular	accomplishment
11	picked (flowers)	regular	accomplishment
12	*dug (a hole)*	irregular	accomplishment
13	planted (flowers)	regular	accomplishment
14	*ate (cookies)*	irregular	accomplishment
15	*blew (the candles)*	irregular	achievement
16	tied (her shoelaces)	regular	accomplishment
17	lifted (a box)	regular	achievement
18	*gave (his mother a present)*	irregular	achievement

6.5 Results on TEGI (Past regular and irregular)

Again, I start with Chinese data followed by Russian data and results will be presented in both raw numbers and percentages.

6.5.1 Results on TEGI-past for Chinese children

All 34 12-year-olds did the TEGI-past test, but only 32 of the 9-year-olds did the test out of the 39 who participated in the research (two 9-year-olds fewer than in TEGI 3SG). Children were excluded because of time restrictions and/or low proficiency/comprehension problems.

On TEGI Past, the group of older children (12-year-olds) outperformed the group of younger children, correctly marking the verb for past in 66% vs 29% of cases respectively (cf. Table 6.14). For both groups, the percentages of incorrect items (i.e. omission of inflection/bare form) are quite low and the percentage of bare verbs for the older children (8%) is lower than that of younger ones (18%). Again, younger children quite extensively use periphrastic marking at almost a quarter of all their responses. The older children use it to a lesser extent, i.e., in 11% of total utterances.

Table 6.14: Chinese children's performance on TEGI Past

PAST	**Correct**	**Incorrect (Bare)**	**Periphrastic marking**	**Other**	**No response**
9y.o. (n=32)	167/576 **29%**	105/576 **18%**	153/576 **27%**	146/576 25%	5/576 1%
12y.o. (n=34)	401/612 **66%**	46/612 **8%**	69/612 **11%**	96/612 16%	0/612 0%

To calculate accuracy in past, the regular verbs, the correctly supplied irregular verbs, as well as the instances of overregularization (i.e. irregular verbs inflected as regular) were summed up and divided by the number of total test items. As you can see in Table 6.15 below, both groups do better with the regular verbs than with the irregular ones. The proportion of overregularization was calculated against the number of irregular verb contexts only. Overregularization was used slightly more by older children but younger children use it as well. What is important here is the fact that low percentages are attested in children's performance regardless of the verb type (regular- irregular).

Table 6.15: Correct suppliance of past tense markers with regular and irregular verbs and instances of overregularization

	Regular correct	**Irregular Correct**	**Overregularization**
Examples	*The boy jumped.*	*The boy ate cookies.*	*The girl digged a hole.*
9y.o.	90/320 **28%**	41/256 **16%**	36/256 **14%**
12y.o.	218/340 **64%**	131/272 **48%**	52/272 **19%**

Let us now see what happens within regular items; do Chinese children produce forms differently depending on whether past is marked in a syllabic way? Table 6.16 below shows the raw sum of accurate answers from all Chinese learners per item. Items 1, 13, and 17 (i.e.

painted, planted, lifted) presented in grey colour have syllabic past tense marking in contrast to the rest where the past tense marking is not syllabic (e.g. cleaned, kicked). It appears that only item 17 may be more problematic, that is, the verb 'lift'.

Table 6.16: Raw sum and percentages of accurate answers in TEGI past per (regular) item.

CH	Item	Sum of accurate answers	Total Regular items	Percentage of mean accuracy/item
1	1	29	66	44%
2	4	31	66	47%
3	5	30	66	45.5%
4	6	31	66	47%
5	8	34	66	51.5%
6	9	35	66	53%
7	11	35	66	53%
8	13	32	66	48.5%
9	16	32	66	48.5%
10	17	21	66	32%

As in the previous test, '*be + verb(-x)*' was found to be the most common pattern of the periphrastic marking category and a main pattern in children's responses. Note here that apart from '*be + verb*' or '*be + verb-ing*' another category included the '*be + inflected verb*', as in the example '*She is cleaned the room*'. See Table 6.17 for raw numbers and percentages. The group of younger children use this strategy here as well to a percentage of 23%, accounting for almost a quarter of all their responses, whereas the group of older children produces this structure to a percentage of 11%. Note that for younger children '*be + bare verb*' is the main pattern with the majority of instances (56). Older children use '*be + verb-ing*' more.

Table 6.17: Children's use of '*be + verb'* (including grammatical *'be + verb-ing'*, and ungrammatical *'be + verb'* and '*be + inflected verb'*) in the TEGI Past

	Grammatical + Ungrammatical be + verb(x)	**Total test items**
9y.o.	(47 + 56 + 29) / 576 **23%**	576
12y.o.	(37 + 15 + 16) / 612 **11%**	612

An item analysis shows that the production of the '*be + verb(-x)*' structure is not affected by certain items as all instances are distributed almost equally across verb types-contexts as can be seen in Table 6.18 below, and as was the case with the 3SG.

Table 6.18: Analysis for item effects on the production of '*be + verb(x)*' forms on TEGI Past

be forms	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
9y.o.	6	6	4	5	9	8	6	8	8	9	7	9	6	5	10	9	10	7
12y.o.	4	1	4	5	4	4	5	4	3	3	4	4	4	3	4	3	5	4

9-year-olds used a considerable number of 'other' responses with the main pattern being the use of present participle minus auxiliary (51/576) as in the example '*She kicking a ball'*. Use of present simple (20/576) (e.g. '*He brushes his hair'*) and '*did + verb-ing'* (21/576) (e.g. *He did painting*) were also attested. Finally, there were a few instances of verbs like '*hit'* or '*run'* (some children chose a verb other than the target) that we could not be sure whether they were marked for tense and which were categorised as 'other', along with verbs of a different type requested by the test (e.g. irregular instead of regular) as well as instances of copula '*be'*.

Turning to individual level performance, Table 6.19 shows results for 9-year-olds on TEGI Past (data sorted from the highest score of correct answers to the lowest).

Table 6.19: Individual results for Chinese 9-year-olds on TEGI Past

		Total Past Correct	Total Past Incorrect	Total Periphrastic marking	Other	No response	Total test items
1	CH_GA_24	18	0	0	0	0	18
2	CH_GA_05	17	0	0	1	0	18
3	CH_GA_06	17	1	0	0	0	18
4	CH_GA_19	17	1	0	0	0	18
5	CH_GA_22	17	1	0	0	0	18
6	CH_GA_04	15	1	0	2	0	18
7	CH_GA_10	13	1	0	4	0	18
8	CH_GA_31	12	0	5	1	0	18
9	CH_GA_37	11	1	5	1	0	18
10	CH_GA_18	10	2	0	6	0	18
11	CH_GA_03	9	1	3	5	0	18
12	CH_GA_08	4	8	1	5	0	18
13	CH_GA_28	2	2	14	0	0	18
14	CH_GA_11	1	8	5	4	0	18

15	CH_GA_12	1	4	9	4	0	18
16	CH_GA_21	1	1	12	2	2	18
17	CH_GA_35	1	3	5	8	1	18
18	CH_GA_36	1	8	1	8	0	18
19	CH_GA_01	0	7	3	8	0	18
20	CH_GA_02	0	1	6	11	0	18
21	CH_GA_09	0	2	4	12	0	18
22	CH_GA_13	0	4	0	14	0	18
23	CH_GA_15	0	0	18	0	0	18
24	CH_GA_20	0	3	10	4	1	18
25	CH_GA_23	0	0	15	3	0	18
26	CH_GA_26	0	2	7	9	0	18
27	CH_GA_27	0	4	6	8	0	18
28	CH_GA_30	0	8	2	8	0	18
29	CH_GA_33	0	11	1	5	1	18
30	CH_GA_34	0	18	0	0	0	18
31	CH_GA_38	0	0	11	7	0	18
32	CH_GA_39	0	2	10	6	0	18
	TOTAL	**167**	**105**	**153**	**146**	**5**	**576**

As can be seen from the Table above, there were 5 9-year-olds who correctly marked tense at least 17 times out of the 18 contexts. There were also 14 children who did not produce any correct past tense marking. In addition, 22 children produced some form of periphrastic marking. What this picture tells us is that there are a few Chinese 9-year-old children who seem to have acquired past tense marking but the vast majority have not yet acquired it showing a very low performance and employing other linguistic structures (i.e. '*be + verb(x)*').

As for the 12-year-olds, as presented in Table 6.20 (data sorted from the highest score of correct responses to the lowest), we see that 11 participants correctly marked past tense at least 17 times out of the 18 test items, while 10 scored below 10 (out of 18). Finally, these children also used some form of periphrastic marking with 4 of them producing it to a great extent (see CH_GB_10, CH_GB_03, CH_GB_22, CH_GB_25). All these results show that there is considerable variation; some Chinese 12-year-olds appear to have learnt the past tense marking, some are at an intermediate level producing it more than half of the times, while a few children seem to be at an earlier developmental stage employing other strategies.

Table 6.20: Individual results for 12-year-olds on TEGI past

		Total Past Correct	Total Past Incorrect	Total Periphrastic marking	Other	Total test items
1	CH_GB_16	18	0	0	0	18
2	CH_GB_17	18	0	0	0	18
3	CH_GB_19	18	0	0	0	18
4	CH_GB_24	18	0	0	0	18
5	CH_GB_29	18	0	0	0	18
6	CH_GB_32	18	0	0	0	18
7	CH_GB_04	17	0	0	1	18
8	CH_GB_13	17	1	0	0	18
9	CH_GB_18	17	0	0	1	18
10	CH_GB_21	17	1	0	0	18
11	CH_GB_33	17	0	0	1	18
12	CH_GB_14	16	1	0	1	18
13	CH_GB_20	16	1	0	1	18
14	CH_GB_28	16	0	1	1	18
15	CH_GB_30	16	0	0	2	18
16	CH_GB_06	15	2	0	1	18
17	CH_GB_26	15	2	1	0	18
18	CH_GB_34	15	3	0	0	18
19	CH_GB_15	14	0	4	0	18
20	CH_GB_27	14	4	0	0	18
21	CH_GB_01	13	0	0	5	18
22	CH_GB_09	13	4	0	1	18
23	CH_GB_12	13	0	4	1	18
24	CH_GB_07	10	3	1	4	18
25	CH_GB_23	9	7	2	0	18
26	CH_GB_08	4	4	0	10	18
27	CH_GB_10	3	0	15	0	18
28	CH_GB_11	3	0	3	12	18
29	CH_GB_31	3	8	1	6	18
30	CH_GB_02	0	0	0	18	18
31	CH_GB_03	0	0	17	1	18
32	CH_GB_05	0	0	0	18	18
33	CH_GB_22	0	0	12	6	18
34	CH_GB_25	0	5	8	5	18
	TOTAL	**401**	**46**	**69**	**96**	**612**

To summarize, both groups of Chinese children showed rather low performance on past tense but younger children much lower than older ones. However, especially the group of younger children use periphrastic marking with the most common structure to be the '*be + verb(x)'*. Older children use the structure '*be + verb(x)'* to a lesser extent while they are more accurate and make fewer errors. Finally, there are more children getting the past correct than the present. Let us now see how the Russian children performed.

6.5.2 Results on TEGI-past for Russian children

TEGI-past was also administered to 71 Russian children; 42 12-year-olds and 29 9-year-olds. In the latter group, two children of the 32 who participated in the research did not do this particular task due to time restrictions and problems with understanding related to their low proficiency in English, while one more child was excluded for a similar reason related to poor understanding of the task.

On this test, as is evident from Table 6.21, again older children outperformed the younger children in correctly supplying the past tense (71% versus 23% respectively). Bare forms were rare in both groups of children while there is a difference between the groups as manifested in their percentages of use of periphrastic marking; younger children use this strategy in half of their responses (50%) while older children use it to a much lesser extent (17%).

Table 6.21: Russian children's performance on TEGI Past

PAST	**Correct**	**Incorrect** (Bare)	**Periphrastic marking**	**Other**	**No response**
9y.o. (n= 29)	120/522 **23%**	32/522 **6%**	259/522 **50%**	110/522 21%	1/522 0%
12y.o. (n=42)	536/756 **71%**	39/756 **5%**	126/756 **17%**	55/756 7%	0/756 0%

Before we discuss children's production of periphrastic marking, let us see what the children's performance is on regular and irregular verbs in past tense as well as the overregularizations produced. In Table 6.22, the instances of correctly marked with *-ed* regular verbs or correctly supplied irregular verbs, and cases of overregularizations are provided. Note that the denominators are the obligatory contexts for regular and irregular past tense marking. From these percentages, it can easily be deduced that both groups did better with regular verbs.

Table 6.22: Correct suppliance of regular and irregular verbs in the past condition and instances of overregularization

	Regular correct	**Irregular Correct**	**Overregularization**
9y.o.	63/290 **22%**	22/232 **9%**	35/232 **15%**
12y.o.	298/420 **71%**	135/336 **40%**	103/336 **31%**

Further focusing on learner's performance on regular past tense marking on an item level, Table 6.23 presents the raw sum as well as percentages of accurate answers for each regular verb as produced by all Russian learners. Again, whether a verb is marked for past tense syllabically or not does not seem to play any role apart from item 17 (verb 'lift') which again appears more problematic than all the rest.

Table 6.23: Sum of accurate answers in TEGI past per (regular) item.

RU	Item	Sum of accurate answers	Total Regular items	Percentage of mean accuracy/item
1	1	35	71	49%
2	4	34	71	48%
3	5	38	71	53.5%
4	6	42	71	59%
5	8	35	71	49%
6	9	38	71	53.5%
7	11	40	71	56%
8	13	37	71	52%
9	16	35	71	49%
10	17	27	71	38%

Returning to the issue of the periphrastic marking, following the same rationale as before, the '*be + verb(x)*' was calculated. The grammatical instances of '*be + verb-ing*' and ungrammatical instances such as '*be + verb*', or '*be + inflected verb*' for both groups are given in Table 6.24. The '*be + verb(x)*' structure is the main pattern used by younger children while those older children who used periphrastic marking, also mainly used this pattern. It is important here to note that both groups used '*be + inflected verb'* to a large extent. Examples include '*He is jumped*' instead of '*He jumped*' and '*She is cleaned*' instead of '*She cleaned*'. They also used '*be + verb-ing*' and '*be + verb*'. The percentage of use of this structure is very high showing again that this pattern is robust.

Table 6.24: Children's use of '*be + verb*' (including grammatical '*be + verb-ing'*, and ungrammatical *'be + verb*' and *'be + inflected verb*') in the TEGI Past

	Grammatical + Ungrammatical be + verb(x)	**Total test items**
9y.o.	(86 + 51 + 90) / 522 **43,5%**	522
12y.o.	(23 + 18 + 59) / 756 **13%**	756

Finally, an item effects analysis showed that there are no item effects in the production of this strategy as the number of instances of '*be + verb(x)*' is distributed almost equally across test items. See Table 6.25 in this respect.

Table 6.25: Analysis for item effects on the production of 'be + verb(x)' forms on TEGI Past

be forms	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
9y.o.	12	12	10	14	15	12	12	13	15	14	11	12	17	11	11	13	11	12
12y.o.	6	4	3	7	5	5	5	9	7	8	5	5	5	4	5	7	6	4

To give a little bit more information on children's 'other' responses, these were more common for younger kids than older ones. Thus, while 3 children produced the vast majority of the 'other' answers in the group of older children (43/55), the 'other' answers were more dispersed across younger children as 19 of them produced some form of 'other' responses. For both groups the patterns were the same. These included the progressive participle (e.g. '*She cleaning her room*.') which was the most common 'other' answer, followed by the '*did + verb-ing*' (e.g. '*She did digging*.'). There were only a couple of instances of present simple and use of the copula *be*.

Let us now discuss individual results to shed further light on children's performance on past tense marking and quantify the number of children that mastered it. Considering the 9-year-olds, from Table 6.26 below, only 1 child scored at least 17/18. The vast majority, that is, 20 provided up to 3 correct answers. 23 children also used some form of periphrastic marking showing again that this pattern is well attested. Bare forms were only a few instances given by some children.

Table 6.1: Individual results for 9-year-old Russian children on TEGI past (data sorted from the highest score of correct answers to the lowest)

		Total Past Correct	Total Past Incorrect (Bare)	Total Periphrastic marking	Other	No response	Total test items
1	RU_GA_17	17	0	0	1	0	18
2	RU_GA_15	14	0	0	4	0	18
3	RU_GA_26	12	4	2	0	0	18
4	RU_GA_30	12	2	4	0	0	18
5	RU_GA_21	11	0	7	0	0	18
6	RU_GA_10	10	2	0	6	0	18
7	RU_GA_20	10	1	4	3	0	18
8	RU_GA_23	10	3	1	4	0	18
9	RU_GA_28	8	1	0	9	0	18
10	RU_GA_02	3	2	11	1	1	18
11	RU_GA_04	2	0	12	4	0	18
12	RU_GA_07	2	1	10	5	0	18
13	RU_GA_08	2	0	16	0	0	18
14	RU_GA_18	2	0	16	0	0	18
15	RU_GA_01	1	0	12	5	0	18
16	RU_GA_06	1	7	0	10	0	18
17	RU_GA_09	1	2	5	10	0	18
18	RU_GA_13	1	3	0	14	0	18
19	RU_GA_29	1	0	10	7	0	18
20	RU_GA_03	0	0	7	11	0	18
21	RU_GA_05	0	1	17	0	0	18
22	RU_GA_14	0	0	16	2	0	18
23	RU_GA_16	0	0	18	0	0	18
24	RU_GA_19	0	1	17	0	0	18
25	RU_GA_24	0	0	18	0	0	18
26	RU_GA_25	0	1	17	0	0	18
27	RU_GA_27	0	0	14	4	0	18
28	RU_GA_31	0	1	9	8	0	18
29	RU_GA_32	0	0	16	2	0	18
	TOTAL	**120**	**32**	**259**	**110**	**1**	**522**

From the 12-year-olds, as we can see in Table 6.27, 18 children scored at least 17/18 showing that quite a lot of children may have acquired this feature. The bare forms are very limited and provided by very few children. Finally, 24 children produced at least one instance of periphrastic marking, but the majority of these instances are due to 5 children.

Table 6.27: Individual results for 12-year-old Russian kids on TEGI past (data sorted from largest correct value to the smallest)

		Total Past Correct	Total Past Incorrect	Total Periphrastic marking	Other	Total test items
1	RU_GB_01	18	0	0	0	18
2	RU_GB_04	18	0	0	0	18
3	RU_GB_06	18	0	0	0	18
4	RU_GB_08	18	0	0	0	18
5	RU_GB_12	18	0	0	0	18
6	RU_ GB_18	18	0	0	0	18
7	RU_GB_34	18	0	0	0	18
8	RU_GB_35	18	0	0	0	18
9	RU_GB_02	17	0	0	1	18
10	RU_GB_03	17	1	0	0	18
11	RU_GB_07	17	1	0	0	18
12	RU_GB_14	17	0	0	1	18
13	RU_ GB_26	17	1	0	0	18
14	RU_ GB_28	17	0	0	1	18
15	RU_GB_33	17	0	1	0	18
16	RU_GB_36	17	1	0	0	18
17	RU_GB_38	17	0	1	0	18
18	RU_GB_42	17	0	1	0	18
19	RU_GB_10	16	2	0	0	18
20	RU_GB_15	16	0	2	0	18
21	RU_ GB_27	16	1	1	0	18
22	RU_GB_32	16	0	2	0	18
23	RU_GB_41	16	0	2	0	18
24	RU_GB_05	15	0	2	1	18
25	RU_ GB_21	15	3	0	0	18
26	RU_ GB_29	15	1	1	1	18
27	RU_ GB_20	14	1	3	0	18
28	RU_GB_31	14	1	3	0	18
29	RU_ GB_22	13	2	3	0	18
30	RU_ GB_25	13	2	2	1	18
31	RU_GB_40	12	3	2	1	18
32	RU_GB_16	10	1	6	1	18
33	RU_ GB_30	9	1	8	0	18
34	RU_GB_09	4	0	14	0	18
35	RU_GB_39	3	11	1	3	18
36	RU_GB_11	2	0	16	0	18

37	RU_GB_37	2	0	16	0	18
38	RU_GB_23	1	1	0	16	18
39	RU_GB_13	0	2	5	11	18
40	RU_GB_17	0	2	16	0	18
41	RU _GB_19	0	0	17	1	18
42	RU_ GB_24	0	1	1	16	18
	TOTAL	**536**	**39**	**126**	**55**	**756**

Overall, results for the Russian cohort showed that older children outperformed the younger ones. There were many 12-year-olds who seem to have acquired the past tense marking. However, 9-year-olds used past tense marking to a more limited extent while they overused '*be + verb(x)*'.

6.5.3 Summary

So far, we saw that on TEGI-Past, both Chinese and Russian 12-year-olds did better than 9-year-olds as in the case of 3SG-agreement.

Group results of older learners showed that Russians seem to do slightly better than Chinese (12y.o.: 71% vs 66%) but this needs to be tested statistically later on. In addition, individual results show that several of the Russian 12-year-olds (N= 18/42) consistently mark verbs for past tense (marking at least 17 out of the 18 verbs) while a considerable number of Chinese (N=11/34) counterparts do as well. As for younger children, Chinese seem to produce more past tense forms (CH: 29% vs RU: 23%), a result to also be tested statistically.

9-year-olds also use '*be + inflected verb*' which was not used in the previous test (TEGI 3SG). However, in the Russian groups '*be + inflected verb*' is more common than '*be + bare verb*'. For the Chinese children, the latter pattern characterizes the performance of the 9-year-olds more, while for those Chinese 12-year-olds who used such a pattern, this was mainly the '*be + verb-ing*'. This seems to indicate an L1 difference between the two groups that we need to explain.

6.6 Summary of the main results of the chapter

Let us now summarise the main findings from the analyses of the TEGI tasks:

Main findings regarding morphology:

- there is low accuracy across ages and L1s

- inflectional morphology is a problematic domain as no group exceeded 71% of accuracy and considering the usually assumed traditional thresholds of 80% (Andersen, 1978; Jia, 2003) or the more conventional (Paradis et al., 2016) 90% (Brown, 1973), it seems that children have not fully acquired the morphemes showing persistent difficulties

Main age effects:

- older children consistently outperform the younger ones
- roughly speaking – there was an increase of around 35% in percentages of accuracy from younger to older ones (e.g. from 10% to 40%)
- improvement is visible – but it does not mean that older children have acquired the features
- the periphrastic marking was attested across L1s – here an age effect is found as older children use the structure less, which is instead more dominant in the younger children with the exception of Russian 12-year-old children on 3SG-agreement
- bare forms also decrease with age – another age effect

Main L1 effects:

- less bare forms in Russian than Chinese
- there is no improvement in percentage of bare forms in Chinese 12-year-olds in 3SG
- the periphrastic marking persists in older Russians while the bare forms persist in older Chinese
- the periphrastic structure '*be + bare verb*' is the main tendency in Chinese children while the '*be + inflected for past form*' is the main pattern used by Russian children in TEGI past

Morpheme effects:

- Asymmetry in the acquisition of 3SG-agreement vs past; children were much more accurate in past than agreement in all age groups and across L1s
- Whether past tense was marked syllabically or not did not seem to play any role in accuracy

Individual variation:

- a lot of individual variation within groups

- a few children performed at ceiling in 3SG; 4/34 Chinese 12-year-olds scored 90% or above on TEGI 3SG and 5/42 Russian counterparts
- more children performed at ceiling in TEGI past
- a few children resort only to periphrastic marking (e.g. 5 Chinese 9-year-olds in 3SG and 3 Russian 9-year-olds use 9/10 or 10/10 periphrastic marking instead of bound morphology)

In the next chapter, I present results on narrative production to see whether children's performance further confirms or weakens the findings presented in this chapter before statistical modelling and discussion take place in the chapters that follow.

CHAPTER 7

RESULTS ON FINITENESS MARKING IN NARRATIVE PRODUCTION

7.0 Introduction

In the previous chapter, I presented results on finiteness marking in elicited production. In this chapter, I focus on a freer type of elicitation, that is, narratives.

Motivated by the same questions and hypotheses, the goal of the empirical analysis of the narratives is:

- to document finiteness marking in narrative production for various forms and tenses.

Also, based on the findings reported in Chapter 6, I will further explore the periphrastic structure '*is + verb(x)'.* More concretely, I will further investigate

- the distribution of different forms (e.g. '*be + stem verb'*, '*be + verb marked for past'*)
- the properties of the periphrastic marking:
    - whether the periphrastic structure agrees with the subject preceding in person and number so as to establish if this construction is actually an *'aux + verb'* construction or whether '*is'* is a frozen item,
    - whether *'was'/'were'* are used bearing the feature of past tense, and
    - to examine the lexical aspect of verbs (e.g. statives, activities) used in the periphrastic structure so as to evaluate the hypothesis that this structure marks aspect rather than tense.

Note though that the periphrastic marking analysis will be presented in Chapter 8. I decided to dedicate a separate chapter in this structure as it is a very interesting pattern not found in the input and as such it could perhaps illuminate acquisition processes. This organization will allow me bring results from both TEGI tasks and narratives together to get a comprehensive picture of its characteristics.

In what follows, I first present in section 7.1 the coding schema I used to annotate, organize and score the relevant properties in narratives. In section 7.2, I describe the annotation tool I used. I then discuss some additional measures of proficiency in section 7.3 and present a comparison between retelling and telling modes in section 7.4. Finally, I present the results in section 7.5 followed by a summary of the main findings in section 7.6.

7.1 Coding and annotation of the narratives

I coded and scored the narratives following three steps: 1. Verb identification, 2. Verb annotation, 3. Organisation of the data and scoring, which I describe in detail below.

Step 1: Verb Identification

First, I identified all the verbs that appeared in each story. Consider an example story below with the verbs in bold.

Example of a (retelling mode) story as produced by a Chinese 9-year-old:

*'The mouse **is** under the tree and the dog **look** the mouse and the dog **want** to catch the mouse. The mouse **is** behind the tree and dog **is run**. A boy is ..., a boy **is going**. He **has** a yellow balloon and a bag of ..., a bag of hotdog. And the mouse **is** small, he run, it **runs** very fast. The dog is ..., the dog **is** down the tree and it **is** very angry. The boy's balloon **is fly** to the tree. He **goes** to the tree and he is ..., he **is want** to catch the balloon and the dog **want** to eat the hotdog. The boy **is jump** and **catch** the balloon. The dog **is put** the bag down. And the boy **is** happy because he **catch** the balloon. The dog **is** happy, too, because he **is eat** the hotdog.'*

(Number of verbs identified: 22)

Overall, I encountered 4,330 verb-tokens at first place considering all stories by all children.

Verb identification: special cases

In general, verb identification was straightforward except when the intelligibility of the verbs was in question.

Thus, whenever a verb form was unintelligible, it was excluded from the analysis and was not coded. For example, when a child said something like '*the dog (tr....) the cat*' or '*He is /googling/ to get his ball back*' or '*the dog was /jagging/ the meat*' that I was not sure what the main verb was, it was left out. These constituted 48 instances in total across all children and stories.

Cases of repetition of the same verb and contracted forms were not straightforward:

In case of repetitions as in '*He, he get, then he get his red ball back*' or the '*dog is, the dog is down the tree'* I scored only one of them, normally the verb with the complement, in this case the second one. I found a total of 115 instances of repetitions, double or multiple.

In case the child attempted to correct him/her-self, I ignored the first instance and did not code it, with the second one being scored only. There were 121 such cases. For example, in an utterance such as '*he run, it runs very fast*' or '*a dog sees, see the cat.*' I always scored the second (or last) instance irrespective of accuracy. Similar coding was followed by e.g. Jia & Fuse (2007) who ignored the first instances and coded the follow-up spontaneous self-corrections.

Finally, contracted forms such as '*It's in the lake'* that only involved present were scored separately for agreement (88 cases in total).

Step 2: Verb coding/annotation

At the second step, I coded and annotated verbs for type, accuracy, and lexical aspect. Table 7.1 shows the coding I followed for the first two categories (i.e. verb types, and accuracy) along with examples. Information with respect to lexical aspect is included in Table 7.4 later in the section.

Table 7.1: Coding schema in narratives; data annotated for verb type, accuracy, and type of error

CODING		
Verb types	Accuracy	Examples
Copula Be	Correct (inflected)	*They are happy.*
	Incorrect (error: agreement)	*The mother and brother is afraid.*
	Incorrect (error: omission)	*The birds happy.*
Auxiliary be (in progressive context)	Correct (in progressive cont.)	*The boy is running.*
	Incorrect (error: agreement)	*The boy are running.*
	Incorrect (error: omission)	*The boy running.*
Other auxiliaries	Correct	*The boy doesn't know ...*
	Incorrect (agreement error)	*and he don't notice that ...*
	Incorrect (error: omission)	*... and his sister not see her ...*
Main thematic verb	Correct	*The boy runs.* *The boy used the fishing rod.* *The dog caught the mouse.*
	Incorrect (error: bare)	*The dog look the mouse.* *... and he make the first small sheep go on the land.*
	Incorrect (error: periphrastic)[19]	*He is want to catch the balloon.*
	Incorrect (other)	*The mother to catching the food*
	Unscorable	*- verbs like hit, run, put*[20] *- unintelligible/ambiguous verbs*
Modal verb (can, will, could, would)	-	*Next, the bird mom will give the little birds ...*

[19] I coded periphrastic structures as those structures where '*be'* is used together with a main verb being it bare or inflected for 3SG or past tense. In other words, in cases of '*be'* overgeneration. In TEGI, I also coded for instances of periphrastic marking which also accounted for erroneous use of present progressive in habitual contexts. Here, the category 'periphrastic marking' does not include instances of present progressive which are counted separately. However, whenever comparisons are to be made only the 'be + verb(-x)' will be taken into account and not the progressive structure.

[20] More info on these verbs are given below in this Section: 'Verb accuracy: Special cases'.

Let us now see the information presented in the table in some more detail.

Verb types

As can be seen, I annotated verbs as: copula *be,* auxiliary *be* (always in progressive context, that is, when preceding the progressive participle), main thematic verbs, other auxiliaries such as *do* and *have*, and modal verbs (i.e. *will, can, would, could*). Note here that auxiliary *be* when preceding a past participle as in '*the birds are saved*' (10 cases) was categorised with 'other auxiliaries' such as *do* and *have*.

Verb accuracy & Error type

Turning to accuracy, one important question is how to define and measure accuracy. Accuracy in obligatory contexts is a standard measure (since Brown, 1973) that can tell us how often learners provide the correct form when required by the context. Unfortunately, in the current stories, the target tense is often ambiguous (cf. example 1) so it is not possible to determine the obligatory context in a reliable way:

(1) *'The dog* ***look*** *at a mouse and the dog* ***want*** *to catch the mouse.'*

Here, for instance, both past tense but also historic present tense would be legitimate. The verb '*look'* could also be used in the present progressive tense as in '*The dog is looking at a mouse ...*' and would not be incorrect. Also, children may use various tenses in one sentence making it difficult to interpret bare forms as in example 2.

(2) '*He* ***wants*** *to get the balloon on the tree and the dog* ***noticed*** *the boy's bag, and he, when the boy* ***gets*** *the balloon, the dog* ***eat*** *the sausage and he* ***was*** *very happy.*'

Given the ambiguity of the contexts and the indeterminacy of the target form, I decided to use error rate as a measure of accuracy, considering only the immediate syntactic frame. I, therefore, coded correct and incorrect forms as shown in the Table 7.1, without reference to the obligatory context. One consequence of scoring forms without taking context into account, is that correct verb forms were scored as correct even when they were used infelicitously. For example, in a sentence like '*but the boy didn't notice because he is using his fishing bat to get the ball away*' the second verb '*is using'* would be best in past progressive as '*was using*'.

I thus annotated as correct any main thematic verb form that was correctly inflected in its syntactic frame (e.g. '*the dog saw the sausage*'); when the verb form was bare for agreement or tense (e.g. *and the crow watch them'* or *'(the dog saw the sausage) and go'*) as well as when the child produced a periphrastic structure such as '*dog is run*' I annotated them as incorrect. Overregularizations like '*falled'* were also marked as correct.

With respect to copula and auxiliary *be*, I annotated as correct, instances of correct agreement with the preceding subject; omissions of the morphemes as well as wrong agreement were annotated as incorrect.

Similarly, with other auxiliaries (i.e. *do*, *have* and *be* of the passive structure), I annotated agreeing forms as correct and non-agreeing forms and omissions as incorrect. Omission errors were examples such as '*and his sister not see her'* where '*did not*' or '*does not*' would be required.

Finally, I annotated modals (i.e. will, can, could, would) without judging the type of form following them. Note that I also annotated '*will*' as a modal in order to distinguish it from auxiliaries such as *do-does* which are inflected for both tense and agreement.

Following accuracy annotation (i.e. correct/incorrect), I specified the error type annotating for bare forms, periphrastic, other types of errors (e.g. agreement errors), as well as omission of copula, and omission of auxiliary.

Verb accuracy: Special cases

I excluded from the annotation verbs like *run*[21], *put, hit, hurt* since it is unclear if they are used for past tense or are just bare verb forms. In most cases no temporal adverb was present to clarify temporal interpretation. However, I did include instances of these verbs when they were inflected for 3SG e.g. *runs, puts, hits*, or appear in a periphrastic structure, e.g. *is run, is put, is hit.* In total, 208 cases of this type were excluded out of 4,330[22] verb-tokens.

A further 125 verb-tokens were excluded from analyses. There were a variety of issues in each case, e.g. two main verbs in a sentence, verbs following 1st singular subjects or 3rd plural subjects. For details on the excluded cases, see Appendix D.

To summarise -as Table 7.2 shows- through scoring I found a total of 4,330 verb-tokens of which 384 verb forms were excluded, leaving 3,866 verbs for analysis. Note that modals were not excluded but not added either to the verbs for further analysis.

---

[21] Especially for *run* it was difficult to perceive whether children uttered run as /rʌn/ and ran as /ræn/. Primarily, I do not even know if children know/make this distinction.

[22] This is the total number of verb tokens before any exclusions.

Table 7.2: All verbs; included and excluded from scoring forms

	Raw Number
Unintelligible verbs/Ambiguous cases	48
Verbs such as *hit, put, run*	208
Other excluded cases	125
Modals	83
Verbs included in the study	3,866
TOTAL NUMBER OF VERBS	4,330

Total number of verbs identified, scored and included in the analysis

Table 7.3 presents the total raw numbers of use of the verbs of all children and all stories for each verb category after the exclusion of unscorable items (see Table 7.2 above). I noted omissions in brackets; remember that these were scored as incorrect instances as they were obligatory in their syntactic frame as in the example '*some (goats) (are) drinking water*' where elements in brackets were missing. Omissions then although in the case of copula and auxiliary *be* are not actually uttered, they are calculated as incorrect answers as they are required by the context.

To calculate the total number of verbs, I added up the total instances of copula '*be'* (plus the omissions which were taken to be incorrect instances), the other auxiliaries (i.e. '*do'*, '*have'*, and the auxiliary '*be'* preceding the past participle), the present progressive ('*be* + *verb-ing'*), and the main thematic verbs inflected, bare, or periphrastic. Modals which were 83 instances in total were not included in this calculation.

Table 7.3: Total number of verbs of each verb category after all exclusions

	Raw sum number of 'obligatory' uses in syntactic frame
Copula *be*	669 (+36 omissions) Total= **705**
Auxiliary *be* -preceding the progressive participle (*aux be + verb-ing*)	370 (+ 115 omissions) Total= **485**
Other auxiliaries: '*be'* preceding a past participle	10
Other auxiliaries: '*do'* & '*have'*	63 (+5 omissions) Total= **68**
Main thematic verbs (inflected for 3SG, past; bare) in 3SG context	**2,322**
Periphrastic marking of main thematic verbs in 3SG and 3PL contexts	273 (3SING) + 3 (3PL) = **276**
ALL VERBS (copula *be* + other auxiliaries (*do, have, be + past participle*) + all main thematic verbs)	**3,866**

Verb annotation for lexical aspect

I annotated verbs for lexical aspectual class following Vendler (1967) and the scoring method provided by Shirai & Andersen (1995). Table 7.4 shows the classification of verbs as activities, states, accomplishments or achievements as proposed by Vendler (1967) along with relevant examples from the present study.

Table 7.4: Classification of verbs according to the lexical aspectual class and examples

Lexical aspectual classes	**Example verbs**
state	*want, enjoy, know*
activity	*run, fly, drink*
accomplishment	*climb the tree, eat the sausage, fly away*
achievement	*catch, see, lose*

To reach a decision about the lexical aspect of a verb, I considered not just the verb but also the arguments or adjuncts following the verb. For example, '*eat'* is coded as an activity, but '*eat a sausage*' as an accomplishment. Note that the definiteness of the argument can influence lexical aspect: '*eat sausages'* may well be an activity while '*eat the sausages*' is an accomplishment.

I excluded cases where it was not possible to determine lexical aspect reliably. For example, '*lead'* as in '*and led the young sheep go back ...'* or '*pick up'* as in *'He pick up a cat to the run'* were excluded. In case of '*lead'*, its meaning in this context is ambiguous -perhaps a more felicitous verb could be '*forced'*. As for the verb '*pick up*', again it is not used appropriately semantically, so the lexical aspect cannot be determined. Similarly, I excluded idiosyncratic uses of verbs as *make* in '*the rat make the dog headache'*. A total of 44 cases were excluded this way, some of which were inflected. Note that 44 verbs were excluded for aspect: 38 were main thematic verbs, while 6 cases concerned the verb of the periphrastic structure.

Finally, there were some cases where the child used an idiosyncratic choice of verb to express an event as depicted in the picture prompts. For example, the child said '*bird eat the cat tail*' when the picture shows that the bird bites the cat's tail, making '*bite'* a more felicitous choice. Apart from the picture though, this interpretation is of course also based on our world knowledge expecting a bird to be more likely to bite a cat's tail and not eat it. In such cases, I coded the lexical aspect of the verb used by the child. Thus, I coded '*eat'* as accomplishment, not as achievement which would be the lexical aspect of the more felicitous verb '*bite'*.

Step 3: Organization of the data and scoring

Once I completed the annotation, I exported the annotated files as excel files and organised the data into columns and categories. For each child, I calculated the proportion of main thematic verbs that were correct (inflected for present 3SG or past tense but also present progressive) versus incorrect (i.e. bare, periphrastic, other). I did the same for all the other verb categories according to the coding schema as appeared in the Table 7.1 above.

Periphrastic structures were counted, and they were then analysed in three ways. First, I counted the frequency and forms used in the periphrastic structure (i.e. '*be + stem verb*' as in e.g. '*the boy's ball <u>is fall</u> down on his hand'*, '*be + verb inflected for 3SG'* as in e.g. '*it'<u>s helps</u> the cat to fall on the ground'*, '*be + inflected for past verb*' as in e.g. '*At this time, the dog <u>was saw'</u>*. Note that some of the cases included in our category of periphrastic included potential passives, i.e. *'is caught, is used, etc.'* where *be* is followed by a past participle. However, the context clarified that the verbs were used actively with both intransitive and transitive verbs followed by an object indicating use of active voice rather than passive, e.g. '*The boy is used the fishing rod*', or '*the butterfly is flew away*'. As mentioned above, there were only 10 instances of '*be + past participle*' such as '*it's caught in the tree branches'* or '*the dog is broken'*, where the construction did indeed indicate a passive use. I excluded these instances of '*be + past participle*' from the analysis of the periphrastic structure and considered the passive auxiliary together with other auxiliaries as described earlier in the section (cf. 'Verb accuracy & Error type' section).

Furthermore, instances such as '*the dog is caught the mouse'* indicate that the periphrastic structure cannot be analysed simply as an incorrect present progressive but needs further analysis. There were 38 instances of '*be + verb inflected for past'* and 9 cases of '*be + inflected for 3SG verb*' out of the total number of periphrastic structure instances (276) which also included '*be + bare verb'*.

Table 7.5 presents the coding of the periphrastic structure that aims to identify all the different forms of it produced by the child L2 learners.

Table 7.5: Classification of forms of '*be + verb(-x)*'

BE + VERB(x): forms	
be + stem verb	*The boy is run.* (229 instances)
be + inflected for 3SG verb	*The boy is runs.* (9 instances)
be + inflected for past verb	*The boy is used the…* *The dog is caught the mouse.* (38 instances)
Total number of periphrastic structures	276

Then, in order to find out whether the verb bears features of agreement, number, and tense, periphrastic structures were scored accordingly.

Table 7.6 shows evaluation of auxiliary *'be'* for agreement and number features; being correct when agreeing with the preceding subject and being incorrect when it did not.

Table 7.6: '*be + verb(-x)*' coded for agreement and number features

BE (+ VERB(x)): agreement			
Agreement	Subject	Form of the periphrastic structure	Example
Correct	3SG	is/was + verb(-x)	*'The boy is run.'*
	3PL	are/were + verb(x)	*'But mom and little goat are drink water in the lake.'*
Incorrect	3SG	are/were + verb(x)	*'... and cat are run out.'*
	3PL	is/was + verb(-x)	*'The mother goat and baby goat is run.'*

Table 7.7 shows results on evaluation of '*be'* of the periphrastic structure for tense features. I calculated the number of instances of '*is'*/'*are'*, and '*was'*/ '*were'* while contracted forms were counted separately.

Table 7.7: '*be + verb(-x)*' coded for tense

BE (+ VERB(x)): tense		
instances of 'is'/'are' (present)	is + verb(x) are + verb(-x)	*But butterfly is fly away.* *His ball are fall to lake.*
instances of 'was'/'were' (past)	was + verb(-x) were + verb(-x)	*The cat was see a yellow butterfly.* *Cat were g(r)o(w) up to the tree.*
contracted be	(subject)'s + verb(-x)	*It's look at the parents ...*

7.2 Annotation tool

To annotate the data of the narratives, I used the WebAnno (Eckart de Castilho, Mújdricza-Maydt, Yimam, Hartmann, Gurevych, Frank, & Biemann, 2016), a web-based annotation tool which supports morphological, semantic and syntactic annotations. As I looked specifically at verb morphology, I created annotation layers relevant to my design and my research. These layers included the verb type (main verb, copula be, auxiliary be, other auxiliary, modal verb, progressive participle, past participle), accuracy (correct, incorrect), error type (bare, omission of auxiliary be, omission of copula be, periphrastic marking, other), and Aktionsart (stative, activity, accomplishment, achievement).

7.3 Additional measures of proficiency

Narratives provided fertile ground to further assess children's proficiency – in addition to the Renfrew Word Finding Vocabulary task, an elicited production task with restricted choices. These measures which will provide us with a more global view of children's performance included 1) Narrative length, 2) Lexical Diversity, and 3) Syntactic Complexity. They were run on both retelling and telling modes, and will be described in detail below.

To obtain scores for these measures, I used the Common Text Analysis Platform (CTAP), a web-based tool for automatic complexity analysis (Chen & Meurers, 2016).

To import data in the CTAP, I first cleaned the data from false starts, repetitions, and 'restructurings' or attempts for 'self-correction' which in any case I had excluded from relevant coding. In some cases, I added punctuation to help the system work better. To give an example of a story before and after 'cleaning', see below:

Example story of a Russian 12-year-old (retelling mode) before:

'*So, the cat saw a beautiful yellow butterfly and wanted to catch that. He jumped but the butterfly flew away. Then, the boy with the, with bucket of fish saw how cat /felt/ in the bush. But when he saw it, he was as scary that his ball jumped away. So, the cat in the bush was very angry but then it saw a lot of fish in the orange bucket. So, he wen.., went out of the bush and start eating that fish. The boy with the /rat/ catch his ball and he was as happy as he can because his ball didn't swam away. And, at the last picture, cat eat all fish so it's very bl.., hap.., happy because he eat all fish.'*

Example story of a Russian 12-year-old (retelling mode) after:

'So, the cat saw a beautiful yellow butterfly and wanted to catch that. He jumped but the butterfly flew away. Then, the boy with bucket of fish saw how cat felt in the bush. But when he saw it, he was as scary that his ball jumped away. So, the cat in the bush was very angry but then it saw a lot of fish in the orange bucket. So, he went out of the bush and start eating that fish. The boy with the rat catch his ball and he was as happy as he can because his ball didn't swam away. And, at the last picture, cat eat all fish so it's very happy because he eat all fish.'

I then ran the analysis in CTAP and got the results in excel files which I further processed statistically. In order to find out whether there were significant differences among the groups with respect to the three additional measures, I carried out three separate Analyses of Variance (one-way independent-measures ANOVA) for each baseline measure in the R software (R Core

Team, 2014). As a first step, I checked whether ANOVA's assumptions were met (1. Independent observations, 2. Outliers, 3. Equality of variances, and 4. Normality of the data). I used data visualization to check for outliers and whenever needed outliers were removed in order not to affect ANOVA's robustness. I tested for equality of variances through the means of Levene's test; in case of a non-significant outcome, this implies that the assumption of equality of variance among the groups is achieved while in case of a significant result, equality cannot be assumed and a different test needs to be used. Finally, with respect to the normality of the data assumption, the group sample sizes were all above 30 and according to the central limit theorem the sampling distribution tends to be normal if the sample is large enough (n>30). Subsequently, I ran the ANOVAs using the aov[23] function. In case of significant results, I ran a post-hoc test to see which group(s) differ using the Bonferroni method in case the assumption of equality of variances was satisfied. In case the ANOVA's assumptions were not met, an alternative to ANOVA was used.

7.3.1 Narrative length: number of clauses

To measure narrative length, the number of clauses were automatically calculated by CTAP. Figure 7.1 shows the average number of clauses for all four groups (CH_9, CH_12, RU_9, RU_12). The averages were similar, apart from the 9-year-olds'.

I then tested results statistically through ANOVA. ANOVA's assumptions were met; (Levene's test: $F(3, 141)=0.1234$, $p= 0.9462$). After the removal of an outlier (from the group of the Russian 12-year-olds), the four groups consisted of 38, 34, 32, and 41 children respectively.

[23] aov() function is an object performing one-way ANOVA in R.

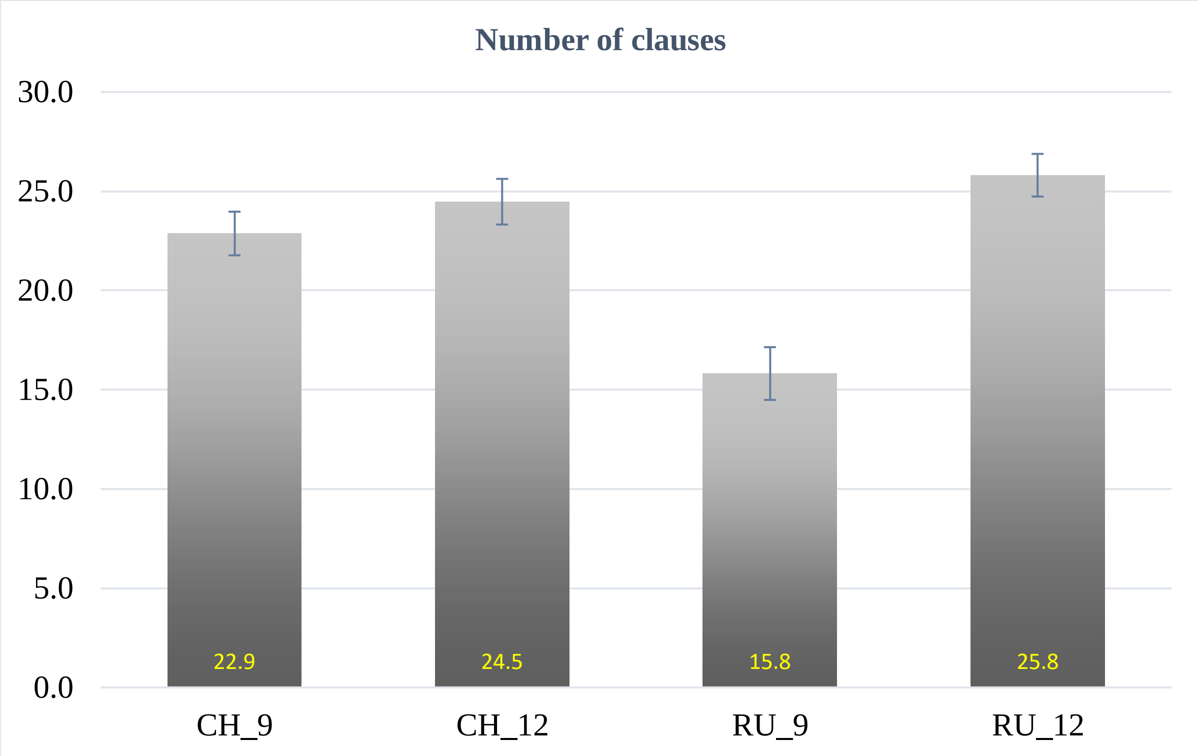


Figure 7.1: Raw average number of clauses produced by the four groups and standard errors

The main analysis with number of clauses as dependent variable and age and language group as independent variable showed that there was a significant difference between the group means, $F(3, 141)=14.46$, $p<0.001$. Post-hoc test using the Bonferroni method revealed that the Russian 9-year-old's group differed significantly from all other groups (RU_9 – CH_9: p(0.00015)<.001; RU_9 – CH_12: $p(4.4e\text{-}06)<0.001$; RU_9 – RU_12: $p(3.1e\text{-}08)<0.01$. None other comparison was found to be significant.

7.3.2 Lexical Diversity: Type-Token Ratio

I next employed another index of learner proficiency, i.e., lexical diversity. The specific measure used is the type-token ratio Uber (Dugast, 1989). This feature calculates the Uber index with the formula: TTR = $(\mathrm{LogN})^2/\mathrm{Log}(N/T)$, where T stands for number of word types, N for number of tokens. I chose this measure because it reduces the effect of text length, that is, it is better used when narratives differ in length (normal type-token ratio (TTR) measures are sensitive to text length differences (i.e. the longer the text, the smaller the lexical diversity), (Dewaele & Pavlenko, 2003)). In addition, it is an appropriate measure for early stages of vocabulary acquisition (losing validity from 3,000 words onwards) (Vermeer, 2000) which is the case here with children producing rather short stories.

As Figure 7.2 demonstrates, 9-year-olds do not seem to differ in this measure, while Russian 12-year-olds seem to outperform the Chinese counterparts.

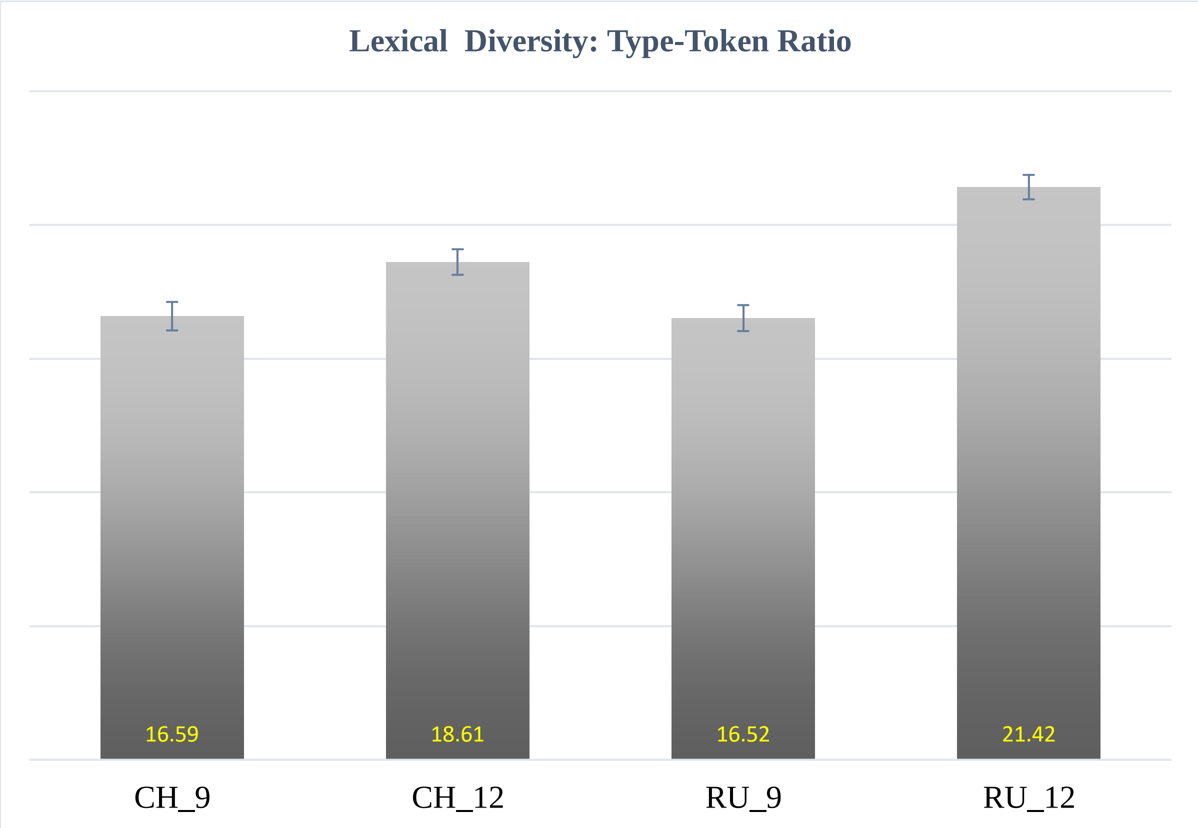


Figure 7.2: Mean type-token ratio values for each group and standard errors

Assumptions were met (Levene's test: $F(3, 141)= 0.891$, $p= 0.4476$)) and numbers in the resulting four groups after the removal of an outlier from the group of Russian 9-year-olds; (CH_9, CH_12, RU_9, RU_12) were 38, 34, 31, and 42 children respectively.

ANOVA analysis with type-token ratio value as the dependent and group as the independent variable showed that there is a significant difference between the group means, $F(3, 141)=23.52$, $p<0.001$. Post hoc tests using the Bonferroni method showed that both older groups performed significantly higher than younger children (CH_12-CH_9: $p=0.02611$; RU_12 – RU_9: $p=6.0e-10$). Chinese 9-year-olds did not differ significantly from Russian 9-year-olds while Russian 12-year-olds performed significantly higher than Chinese counterparts ($p=0.00041$).

7.3.3 Syntactic complexity: number of subordinate clauses

I now turn to a measure of syntactic complexity of the stories by calculating the dependent clause ratio. This measure calculates the number of dependent clauses out of the number of all clauses. As Figure 7.3 shows, the Chinese and Russian 9-year-olds do not seem to differ, however, Russian 12-year-olds seem to use many more complex clauses than any other group.

To confirm this statistically, I checked for ANOVA's assumptions which were partly met (Levene's test result was significant, $F(3, 137)= 4.4999$, $p<0.01$, implying that this assumption was violated). After removing 3 outliers the resulting numbers in each group was 38, 32, 32, 41 respectively.

I then computed Welch's F -an alternative to normal ANOVA appropriate in cases of unequal variances across groups (Field, 2012)- with L1 and age group as the independent variable, and dependent clause ratio as the dependent variable.

This test's result showed a highly significant result, $F(3, 74)=11.157$, $p(4.007e\text{-}06) <.05$. I then used the Games-Howell test, a post hoc test for unequal variances and unequal sample sizes. The Games-Howell test results confirmed that the Russian 12-year-olds' group differed significantly from all the other groups (RU_12-CH_12: $t(67)= 4.44$, $p<.001$; RU_12-CH_9: $t(69)= 5.77$, $p<.001$; RU_9-RU_12: $t(65)=3.87$, $p<.001$) producing a significantly higher number of dependent clauses. No other group pair showed significant differences.

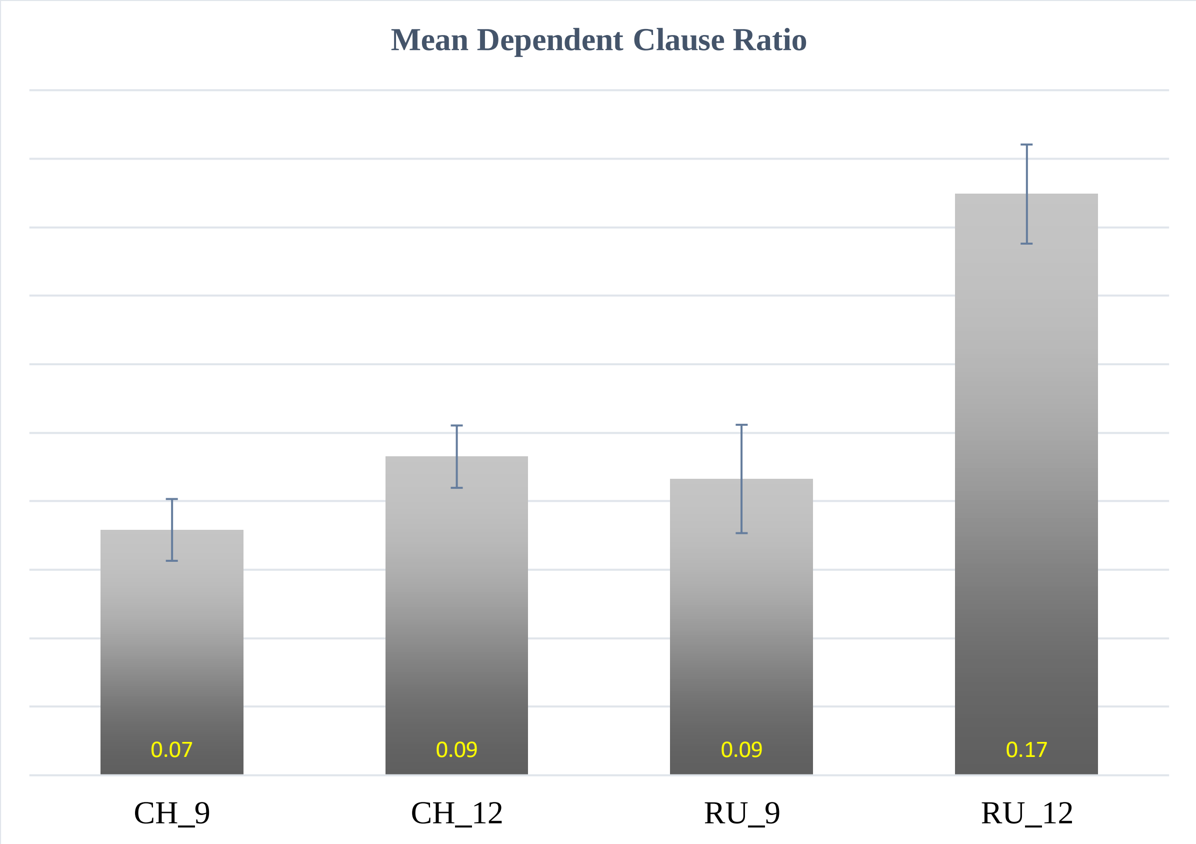


Figure 7.3: Average dependent clause ratio per group and standard errors

7.3.4 Lexical diversity and number of subordinate clauses; is there an effect of vocabulary for older Russians' higher performance?

In order to address whether 12-year-old Russian learners have overall larger lexical diversity and use a higher number of complex clauses than Chinese counterparts due to their better vocabulary scores, I ran two multiple regression analyses in the R software (R Core Team, 2014) with either lexical diversity or ratio of subordinate clauses as the dependent variable and the interaction of Renfrew Vocabulary score with nationality group (CH or RU) as independent variables. This interaction tests whether the effect of the Renfrew vocabulary score differs across the two nationality groups and whether such an interaction predicts lexical diversity.

The results of the first analysis where lexical diversity was the dependent variable showed only two significant main effects; vocabulary and L1. Vocabulary was a highly significant predictor of lexical diversity ($t = 3.924$, $p<0.001$); higher vocabulary scores meant increasing lexical diversity. L1 also was a significant predictor with Russians outperforming Chinese learners ($t = 2.947$, $p<0.01$). The interaction between vocabulary and L1 was not significant meaning that Russians did not have higher lexical diversity than Chinese because of higher vocabulary scores.

The adjusted R^2 value of the model was 0.3911 so 39% of the variation in lexical diversity can be explained by the model containing vocabulary scores and learners' L1. The data met the assumptions of data linearity, homoscedasticity, no multicollinearity and normality of residuals.

Turning to the second analysis, where ratio of subordinate clauses was the dependent variable with the same independent structure as described above, it was shown that there was only one significant main effect, that of L1. L1 could predict the dependent clause ratio ($t=5.285$, $p<0.001$) with Russian L1 predicting higher number of complex (subordinate) clauses than Chinese.

The adjusted R^2 value of the model was 0.29 so 29% of the variation in dependent clause ratio can be explained by the learners' L1. The data met the assumptions of data linearity, homoscedasticity, no multicollinearity and normality of residuals.

Taken these results together, the fact that Russian 12-year-olds have larger lexical diversity and use more complex clauses than Chinese counterparts cannot be attributed to their higher vocabulary scores/proficiency and perhaps linguistic distance is strengthened as a potential explanation for their performance.

7.3.5 Summary

Considering the results of all tests and measures used to assess proficiency, we saw the following:

- *Renfrew task*: 9-year-olds were not found to differ; Russian 12-year-olds did significantly better than their Chinese counterparts
- *CEFR level of classes*: more Chinese 9-year-olds are in a A2 CEFR level class compared to Russian 9-year-olds – 12-year-olds do not show any difference in this respect
- *Narrative length*: only Russian 9-year-olds differed significantly from all the other groups producing a lower number of clauses
- *Lexical diversity*: Chinese 9-year-olds did not differ from Russian 9-year-olds; Russian 12-year-olds performed significantly better than their Chinese counterparts
- *Syntactic complexity*: only the Russian 12-year-olds differed significantly from all the other groups showing higher syntactic complexity.
- Russian 12-year-olds' higher lexical diversity and number of subordinate clauses is not predicted by their higher scores on Renfrew vocabulary task. Their L1 was a significant predictor in both analyses while Renfrew vocabulary scores had a significant main effect in predicting lexical diversity but for both nationalities.

These results draw an interesting picture with respect to proficiency and the various measures seem to capture different aspects of it.

7.4 Retelling versus telling

Remember that all children produced both a narrative telling and a retelling. I therefore next examined whether mode impacts on verb form accuracy. In this analysis, I included accuracy on all verb forms scored, i.e. the total number of verbs.

Figure 7.4 shows the means of ratios of verb form accuracy in each mode. As can be seen scores are very similar within all groups.

To test whether differences between the modes within each group are statistically significant, I carried out dependent t-test analyses using the R software (R Core Team, 2014) to compare the ratios of the verb form accuracy in the two modes. Assumptions were met; (i) data are measured at least at the interval level, and (ii) normality of the sampling distributions was tested using the Shapiro-Wilks test; for all groups, the results of the Shapiro-Wilks test were non-significant, implying that normality can be assumed.

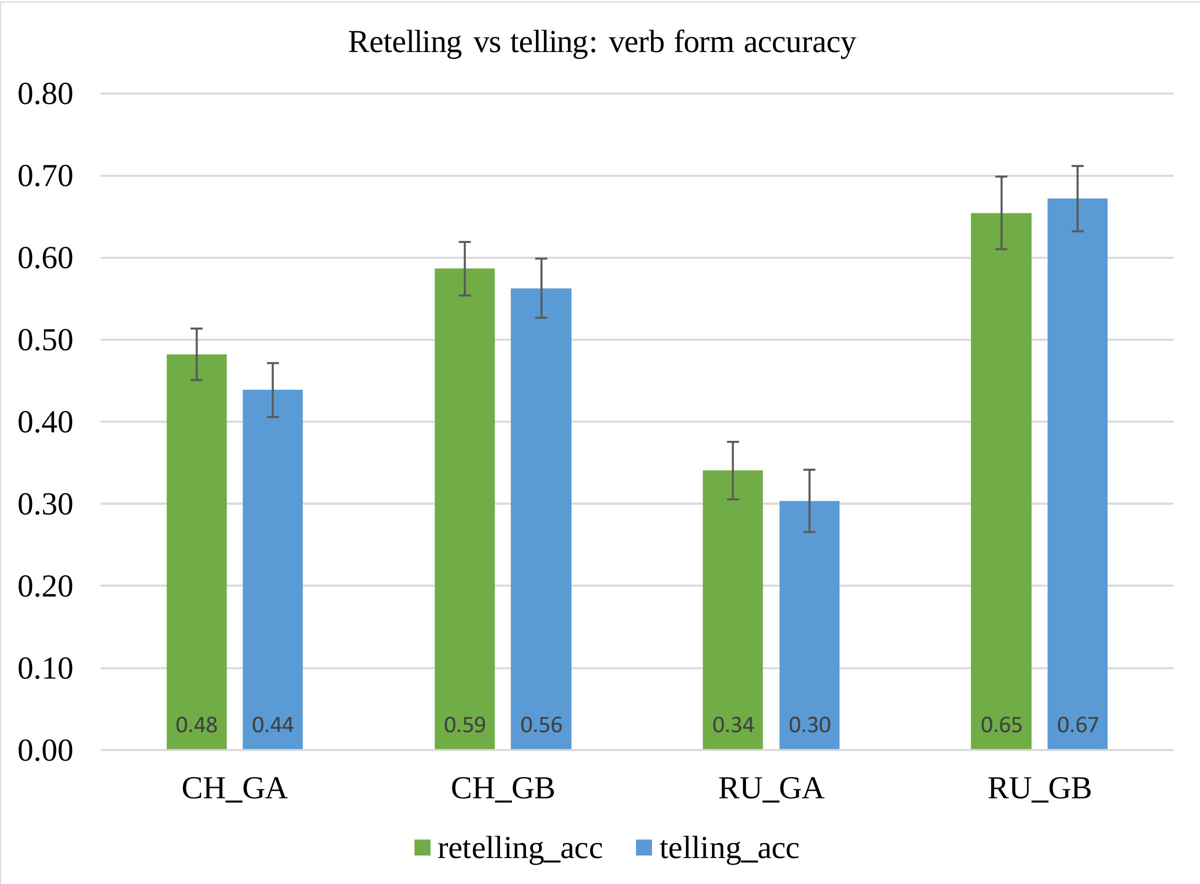


Figure 7.4: Within group-comparisons of verb form accuracy ratios in retelling and telling conditions; means and standard errors

Results of the paired-samples t-test analyses showed that for no group verb form accuracy differed significantly in the retelling and telling modes.

In light of these results, I will not distinguish between the two modes in the analyses presented in the following sections.

7.5 Results on finiteness marking in narratives

I now turn to the results on finiteness marking in narratives. The first question is what children do in terms of the rendition of verb morphology in narrative productions. Specifically, to what extent do children inflect verbs and to what extent do they use bare forms, periphrastic marking, or other forms?

Below I present results for different verbs in the following order: copula *be*, auxiliary *be*, other auxiliaries, and thematic verbs. Results will be presented for each age and L1 group, e.g. Chinese 9-year-olds: CH_9.

7.5.1 Copula be

Table 7.8 presents Chinese children's overall accuracy on copula *be*. The total number of copula *be* produced by Chinese learners is 392(/705) forms. Recall that correct answers involve correctly inflected forms agreeing with the subject, while incorrect answers involve both omissions and incorrect agreement cases (for which I give details below in Table 7.9). As can be seen, the 3SG subjects prevail in both cohorts. Both younger and older groups perform at ceiling on copula *be* in 3SG (CH_9: 99.5%; CH_12: 98%), however, accuracy drops considerably in plural contexts (CH_9: 66.5%; CH_12: 78.5%). Further, the two groups do not seem to differ in the 3SG condition, while older children seem to perform somewhat better in the 3PL condition.

Table 7.8: Chinese children's overall accuracy on copula be in narratives

COP BE		Total Correct	Total Incorrect	Total
CH_9 *n= 38*[24]	3SG	180/181 99.5%	1/181 .05%	181 100%
	3PL	26/39 66.5%	13/39 33.5%	39 100%
CH_12 *n= 34*	3SG	132/135 98%	3/135 2%	135 100%
	3PL	29/37 78.5%	8/37 21.5%	37 100%

Table 7.9 shows scores for use in different tenses, omission, and agreement errors for 3rd person singular and plural. Starting with correct tense use, we can see that both groups generally use more present than past. However, the older children use much more past tense than the younger ones. Specifically, the 9-year-olds use mostly present tense, i.e. '*is*' (77%) while the 12-year-olds use both present and past to a similar extent (48% - 43%). There seems to be a further effect of person, since most past tense use is in 3SG while in 3PL contexts, both younger and older children use mostly the present tense, i.e. '*are*'. Omissions are generally very few for both groups. Agreement errors though are significant, interestingly concentrated in plural, indicating that there seems to be a more general effect of plural in the data. Accuracy improves in the older children in 3SG-agreement errors.

[24] Of the 39 9-year-olds, 1 child could not produce a story and thus is excluded.

Table 7.9: Chinese children's accuracy on copula be in narratives

COP BE		Correct			Incorrect		
		Present	Past	Contracted	Omission	Agreement	Total
CH_9 *n= 38*	3SG	139/181 77%	19/181 10.5%	22/181 12%	1/181 0.5%	0/181 0%	181 100%
	3PL	22/39 56.5%	4/39 10%	0/39 0%	3/39 8%	10/39 25.5%	39 100%
CH_12 *n= 34*	3SG	65/135 48%	58/135 43%	9/135 7%	3/135 2%	0/135 0%	135 100%
	3PL	20/37 54%	9/37 24.5%	0/37 0%	0/37 0%	8/37 21.5%	37 100%

Let us now turn to the Russian children who produced in total 313(/705) copula *be* forms. Table 7.10 shows the Russian L2 learners' overall accuracy with incorrect forms involving both incorrect agreement instances as well as omissions. Again, all children use mostly 3SG subjects. Older children outperform younger ones in accuracy in both 3SG and 3PL conditions; the 12-year-olds have a very high performance on 3SG (92%) while they do quite well although not at ceiling in 3PL (85%). The 9-year-olds seem to have more problems with the copula *be* as in both 3SG and 3PL contexts they have a rather low performance (78% and 62.5% respectively). The number of incorrect answers is also high in case of 9-year-olds. Overall, it seems that there is again an effect of person, while accuracy in both 3SG and 3PL improves with age. Still, the 3PL appears to be a more 'difficult' context for learners.

Table 7.10: Russian children's overall accuracy on copula be in narratives

COP BE		Total Correct	Total Incorrect	Total
RU_9 *n= 32*	3SG	75/96 78%	21/96 22%	96 100%
	3PL	5/8 62.5%	3/8 37.5%	8 100%
RU_12 *n= 42*	3SG	167/182 92%	14/182 8%	182 100%
	3PL	23/27 85%	4/27 15%	27 100%

Let us turn to Table 7.11 which provides information on tense marking, omission, and agreement errors. Starting with correct use of tense forms, we see that like Chinese children, the younger Russian children use mostly present forms. However, the pattern changes with the older children who seem to swift to past. Again, we see an effect of plural; past tense use is exclusively in 3rd singular contexts in the 9-year-olds and significantly higher in 3SG in the older children. Omissions are frequent in 9-year-olds' speech and fewer in older children's. As for agreement errors, these are not very rare although we should also consider the small number of raw occurrences. Again, we see higher error percentages in 3PL.

Table 7.11: Russian children's accuracy on copula be in narratives

COP BE		**Correct**			**Incorrect**		
		Present	Past	Contracted	Omission	Agreement	Total
RU_9 *n= 32*	3SG	48/96 50%	19/96 20%	8/96 8%	15/96 16%	6/96 6 %	96 100%
	3PL	5/8 62.5%	0/8 0%	0/8 0%	2/8 25%	1/8 12.5%	8 100%
RU_12 *n= 42*	3SG	22/182 12%	140/182 77%	5/182 3%	12/182 6%	3/182 2%	182 100%
	3PL	9/27 33%	13/27 48%	1/27 4%	0/27 0%	4/27 15%	27 100%

Copula be Summary

Results on copula *be* revealed the following: (i) accuracy improves in older children for all groups, (ii) past tense use appears later in all groups, (iii) 3PL is a difficult context for all learners, showing higher error rate and lower past tense use, (iv) omission of copula is an issue for Russian children, but not for Chinese. Russian children though do improve with age. (v) There seems to be an L1 difference in the use of past tense for older children; Russian children tell stories predominantly in past while Chinese children are not as productive with past. Finally, (vi) Russian children seem to also be doing better with plural agreement. Overall, it seems that 12-year-old Russians have acquired *be* more than Chinese have.

7.5.2 Auxiliary be

Let us now consider children's performance on auxiliary *be*. Table 7.12 shows accuracy scores for the Chinese groups who produced 210(/485) instances of the progressive structure. Unlike copula *be*, accuracy is not at ceiling. We see again an effect of number, as both groups do much better in the 3SG condition than with 3PL subjects. Of course, the contexts in plural are very few to make any reliable generalisations. Incorrect answers (i.e. incorrect agreement instances as well as omissions) are produced to a considerable extent even in the 3SG condition. (CH_9: 16.5%, CH_12: 17.5%).

Table 7.12: Chinese children's overall accuracy on auxiliary be in narratives

AUX BE		Total Correct	Total Incorrect	Total
CH_9 *n= 38*	3SG	86/103 83.5%	17/103 16.5%	103 100%
	3PL	1/2 50%	1/2 50%	2 100%
CH_12 *n= 34*	3SG	81/98 82.5%	17/98 17.5%	98 100%
	3PL	4/7 57%	3/7 43%	7 100%

Table 7.13 gives a detailed presentation of children's performance on auxiliary *be*. As with copula *be*, both groups use much more the present tense auxiliary than the past in both 3SG and 3PL. Put it differently, although the numbers are only a few for 3PL, past seems to be used exclusively in the 3SG condition. Omissions are not in negligible numbers in the 3SG condition for both groups while agreement errors are generally a few.

Table 7.13: Chinese children’s accuracy on auxiliary be in narratives

AUX BE		Correct			Incorrect		
		Present	Past	Contracted	Omission	Agreement	Total
CH_9 *n= 38*	3SG	76/103 73.5%	8/103 8%	2/103 2%	16/103 15.5%	1/103 1%	103 100%
	3PL	0/2 0%	1/2 50%	0/2 0%	1/2 50%	0/2 0%	2 100%
CH_12 *n= 34*	3SG	54/98 55%	24/98 24.5%	3/98 3%	16/98 16.5%	1/98 1%	98 100%
	3PL	4/7 57%	0/7 0%	0/7 0%	0/7 0%	3/7 43%	7 100%

Consider now Table 7.14 showing accuracy scores of the Russian groups. Russian children, unlike Chinese children, start with low accuracy which improves. Again, there seems to exist an effect of plural; older children perform better in 3SG contexts (82%) than 3PL contexts (61.5%). Incorrect answers are quite high especially for 9-year-olds in both 3SG and 3PL contexts.

Table 7.14: Russian children’s overall accuracy on auxiliary be in narratives

AUX BE		Total Correct	Total Incorrect	Total
RU_9 *n= 32*	3SG	63/123 51%	60/123 49%	123 100%
	3PL	5/9 55.5%	4/9 44.5%	9 100%
RU_12 *n= 42*	3SG	107/130 82%	23/130 18%	130 100%
	3PL	8/13 61.5%	5/13 38.5%	13 100%

Table 7.15 presents detailed information about tense marking, omission and agreement errors. Russian 9-year-olds use almost exclusively present tense in 3SG contexts, while 12-year-olds use past tense to a much greater extent. As for 3PL contexts, younger children exclusively use present while older children use both present and past tense to the same extent.

This again shows an effect of person; there is a clear shift in past tense use which is prominent in 3SG contexts. Omissions are quite frequent, especially for younger children, while agreement errors are few in raw occurrences but not negligible if we consider the percentages.

Table 7.15: Russian children's accuracy on auxiliary be in narratives

AUX BE		**Correct**			**Incorrect**		
		Present	Past	Contracted	Omission	Agreement	Total
RU_9 *n= 32*	3SG	52/123 42%	5/123 4%	6/123 5%	59/123 48%	1/123 1%	123 100%
	3PL	5/9 56%	0/9 0%	0/9 0%	2/9 22%	2/9 22%	9 100%
RU_12 *n= 42*	3SG	36/130 28%	68/130 52%	3/130 2.3%	20/130 15.4%	3/130 2.3%	130 100%
	3PL	4/13 30.8%	4/13 30.8%	0/13 0%	1/13 7.6%	4/13 30.8%	13 100%

Auxiliary be Summary

Analysis of children's production of auxiliary *be* in narratives revealed the following: 1. Lower accuracy in auxiliary *be* compared to copula *be*, 2. Big development in accuracy for Russian but not for Chinese children, 3. More past tense use from Russians, 4. An effect of plural as a context with less past use and more accuracy errors (though numbers are small for definite conclusions), 5. Omission is a bigger problem for Russians than Chinese.

7.5.3 Other auxiliaries

Let us now consider auxiliaries *do, have, be (+ past participle)*. Starting with Chinese groups, they produced 41(/78) instances of such auxiliaries. As Table 7.16 shows, both Chinese groups hardly use any 'other' auxiliaries in 3PL. Use is restricted to 3SG with older children achieving high accuracy (90.5%) outperforming the younger children who have a rather low accuracy (62.5%).

Table 7.16: Chinese children's performance on other auxiliaries in narratives

OTHER AUX		Total Correct	Total Incorrect	Total
CH_9 *n= 38*	3SG	10/16 62.5%	6/16 37.5%	16 100%
	3PL	2/2 100%	0/2 0%	2 100%
CH_12 *n= 34*	3SG	19/21 90.5%	2/21 9.5%	21 100%
	3PL	1/2 50%	1/2 50%	2 100%

Table 7.17 provides a more analytic account of children's performance. As with copula and auxiliary *be*, younger children use more present than past, while the use of past increases in older children. Again, increased accuracy and use of past is restricted to 3SG person, there are very few raw examples of 3rd plural.

Table 7.17: Chinese children's accuracy with other auxiliaries in narratives

OTHER AUX		**Correct**		**Incorrect**		
		Present	Past	Omission	Agreement	Total
CH_9 *n= 38*	3SG	6/16 37.5%	4/16 25%	1/16 6.25%	5/16 31.25%	16 100%
	3PL	2/2 100%	0/2 0%	0/2 0%	0/2 0%	2 100%
CH_12 *n= 34*	3SG	9/21 43%	10/21 47.5%	1/21 4.75%	1/21 7.75%	21 100%
	3PL	1/2 50%	0/2 0%	0/2 0%	1/2 50%	2 100%

Table 7.18 presents the overall accuracy on other auxiliaries for Russian children. Russians produce 37(/78) instances of other auxiliaries. Again, plural subjects are scarce. Younger children show a very low performance with accuracy of 25%. Older ones though seem to perform quite well achieving an accuracy rate of 94%.

Table 7.18: Russian children's accuracy on other auxiliaries in narratives

OTHER AUX		Total Correct	Total Incorrect	Total
RU_9 *n= 32*	3SG	4/16 25%	12/16 75%	16 100%
	3PL	1/2 50%	1/2 50%	2 100%
RU_12 *n= 42*	3SG	15/16 94%	1/16 6%	16 100%
	3PL	3/3 100%	0/3 0%	3 100%

As Table 7.19 shows, younger children use only present tense while older children use mostly past tense. Older children are much more accurate than younger ones who mainly make agreement errors in 3SG contexts such as *s/he don't*.

Table 7.19: Russian children's accuracy with other auxiliaries in narratives

OTHER AUX		**Correct**		**Incorrect**		
		Present	Past	Omission	Agreement	Total
RU_9 *n= 32*	3SG	4/16 25%	0/16 0%	2/16 12.5%	10/16 62.5%	16 100%
	3PL	1/2 50%	0/2 0%	1/2 50%	0/2 0%	2 100%
RU_12 *n= 42*	3SG	4/16 25%	11/16 69%	0/16 0%	1/16 6%	16 100%
	3PL	2/3 67%	1/3 33%	0/3 0%	0/3 0%	3 100%

Other auxiliaries Summary

Considering results on other auxiliaries, we saw the following: 1. Development from younger to older children's performance, 2. Past is used more by older children, 3. Agreement errors are frequent in younger children, 4. Older groups perform similarly; Chinese 9-year-olds outperform their Russian counterparts.

It is important to note at this point that this analysis mainly concerned the auxiliaries and not the forms followed. Thus, there were cases of a correctly inflected auxiliary but if I considered it combined with the following form, the string altogether would be considered ungrammatical *'didn't swam', 'hasn't notice',* etc. Accuracy then concerns only auxiliaries.

7.5.4 Main thematic verbs

Let us finally turn to results from main thematic verbs. In this analysis I will only consider verbs in 3SG contexts as main thematic verbs following plural subjects (N=49) were excluded. Starting with the Chinese groups, they both produced 1,384(/2,595) instances of main thematic verbs. Table 7.20 presents overall overt finiteness marking (i.e. 3SG -*s* and past tense) percentages for the Chinese children. Younger children use 3SG-agreement or past tense inflection at almost a quarter of all their verb production (27.6%) whereas older children mark almost half of the verbs produced for tense and/or 3SG-agreement (47%). The overall overt finiteness marking is quite low for both groups then.

Table 7.20: Chinese children's overall use of bound morphology on main thematic verbs (MTV) in narratives

MTV		Total Correct	Total Incorrect	Total
CH_9 *n= 38*	3SG	187/677 27.6%	492/677 72.4%	677 100%
CH_12 *n= 34*	3SG	334/707 47%	373/707 53%	707 100%

Table 7.21 shows accuracy for the different morphemes, 3SG-agreement and past tense, as well as types of errors. Younger children use both present simple and past tense to a similar extent (13.6% vs 14%) while older children use mostly past tense (17% vs 30%). Omissions are very high in percentages for both groups and do not drop considerably from the younger children to the older ones (55.1% - 48%). Periphrastic marking though drops from 17.3% produced by younger children to 5% in case of older children.

Table 7.21: Chinese children's use of bound morphology on main thematic verbs in narratives

MTV		**Correct**		**Incorrect**		
		Present	Past	Omission	Periphrastic	Total
CH_9 *n= 38*	3SG	92/677 13.6%	95/677 14%	373/677 55.1%	117/677 17.3%	677 100%
CH_12 *n= 34*	3SG	121/707 17%	213/707 30%	339/707 48%	34/707 5%	707 100%

Table 7.22 shows results on overt finiteness marking for Russian learners. Russian children produced in total 1,211(/2,595) main thematic verbs. Younger children mark verbs for tense and/or 3SG-agreement at a percentage of 22.5% of all the main thematic verbs they produce while older children's percentage increases considerably to 59%.

Table 7.22: Russian children's overall use of bound morphology on main thematic verbs in narratives

MTV		Total Correct	Total Incorrect	Total
RU_9 *n= 32*	3SG	99/439 22.5%	340/439 77.5%	439 100%
RU_12 *n= 42*	3SG	456/772 59%	316/772 41%	772 100%

Table 7.23 gives us a detailed picture. Both Russian groups use mostly past tense compared to present. Considering incorrect answers, the main issue at children's performance is the omissions (61%; 34.5% respectively) while younger ones produce a high number of periphrastic marking instances as well (16.5%) which decreases to 6.5% for older children. In other words, accuracy doubles for older children and omission is halved showing a developmental/proficiency effect.

Table 7.23: Russian children's use of bound morphology on main thematic verbs in narratives

MTV		**Correct**		**Incorrect**		
		Present	Past	Omission	Periphrastic	Total
RU_9 *n= 32*	3SG	30/439 7%	69/439 15.5%	268/439 61%	72/439 16.5%	439 100%
RU_12 *n= 42*	3SG	45/772 6%	411/772 53%	266/772 34.5%	50/772 6.5%	772 100%

Main thematic verbs summary

The results on overt finiteness marking showed low percentages of overt finiteness marking on verbs in narratives by all groups and for younger ones even lower than older ones. Younger children marked thematic verbs for finiteness to a very limited extent while older children who outperformed them marked for tense/3SG-agreement in almost half of the verbs used. We see again then a developmental effect with older children being at a more advanced developmental stage than younger children.

While Chinese and Russian 9-year-olds do not seem to differ, 12-year-old Russians achieve a considerably higher percentage (59%) of correct markings than their Chinese counterparts (47%). This reveals an L1 effect.

7.5.5 Interim Summary

Figure 7.5 summarises the results for learners' performance on all morphemes focusing on 3SG contexts. These are the key points:

- all children show the same ranking in the acquisition of morphemes which is as follows: cop *be* > aux *be* > main thematic verb[25]. This shows a developmental progression across morphemes.
- development is very clear considering the Russian groups; in all morphemes, free and bound, older children do much better than younger ones showing a steep increase in accuracy. Considering development for Chinese children, the two groups do not differ as far as copula and auxiliary *be* are concerned, while older ones outperform younger

[25] Note that 'other auxiliaries' are not included here as for younger children their usage in terms of accuracy follows aux *be* while for older children they precede them. However, there are a few instances of 'other auxiliaries' and thus no safe conclusions can be drawn at this point.

ones in correct marking of other auxiliaries and main thematic verbs. The development in case of Chinese children then is seen as flatter than it was for Russians.

- Although Russians start having lower accuracy, in the end they manage to match Chinese children or even surpass them. The lower initial accuracy is probably due to proficiency and educational differences while the steep increase in performance is an L1 effect. Chinese children have persistent problems with main thematic verbs.

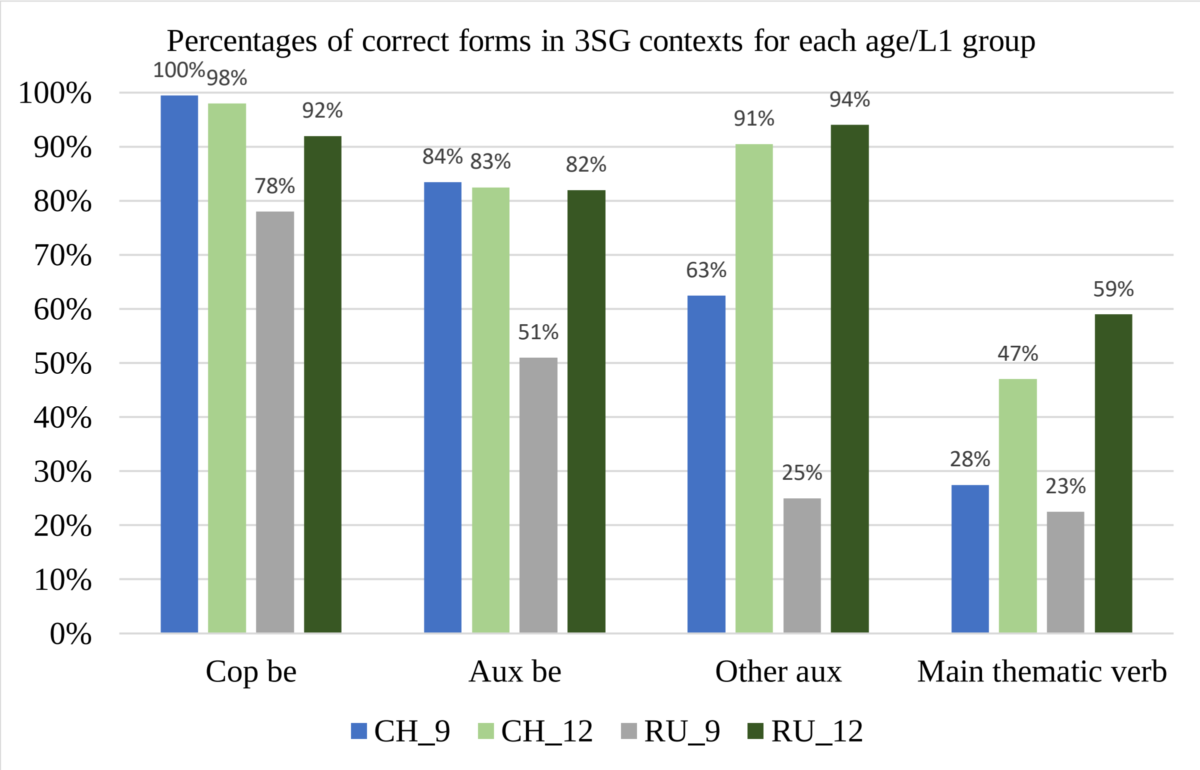


Figure 7.5: Percentages of correct forms following 3SG subjects

Figure 7.6 summarises the results for learners' performance on copula and auxiliary *be* as well as other auxiliaries focusing on 3PL contexts. These are the key points:

- Plural contexts appear much more challenging for copula and auxiliary *be*

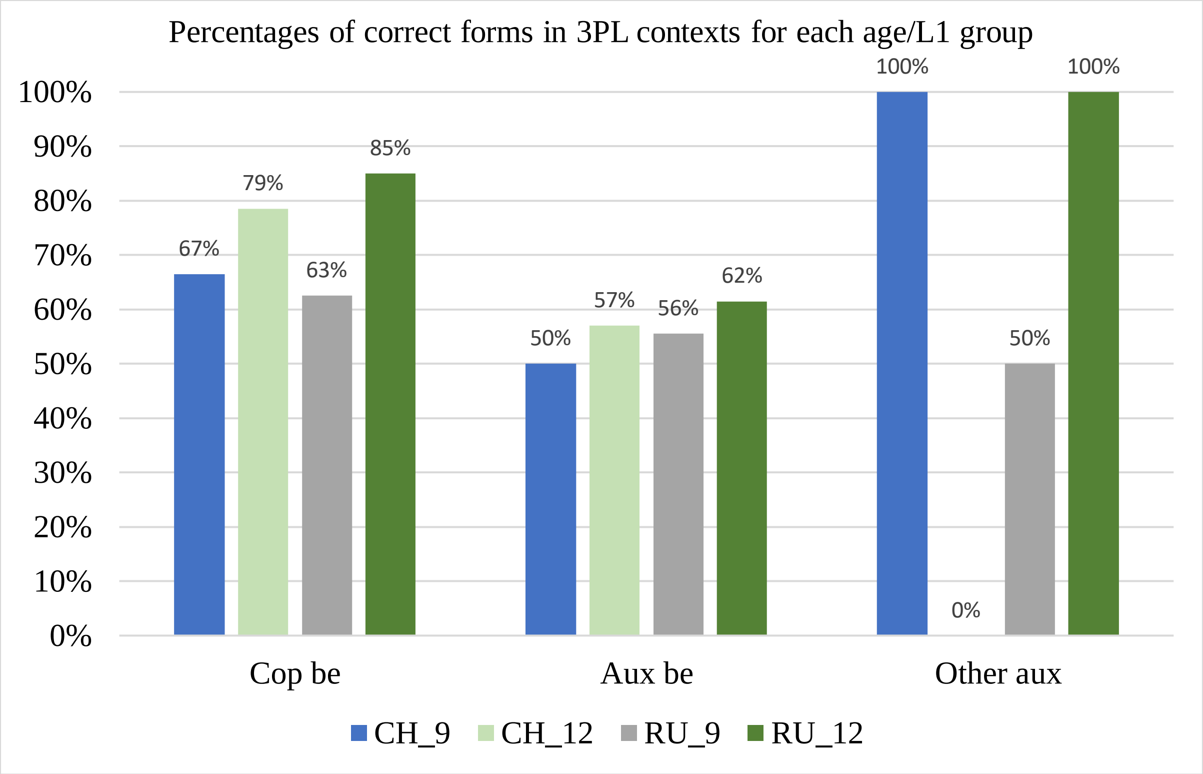


Figure 7.6: Percentages of correct forms following 3PL subjects

To conclude, children seem to have problems with inflectional morphology especially with 3SG-agreement and past tense and even suppletive morphology is not totally unproblematic. Younger children's performance is characterized by omissions but also by the use of the periphrastic structure which is a common finding analyzed further in the next chapter.

Older children outperform the younger ones as in marking of main thematic verbs for tense/agreement and appear to be at a more advanced developmental stage since their performance overall is much more accurate and less diverse in terms of omissions, incorrect agreement instances or periphrastic marking instances.

Considering L1 cohorts, Russian 12-year-olds outperform the Chinese counterparts on overt finiteness marking on main thematic verbs while Chinese 9-year-olds perform similarly to or better than Russian 9-year-olds. Given that proficiency -considering the CEFR level of the classes children attend- between the older groups is similar, then these results show L1 effects. For younger children, the difference attested in performance may be due to proficiency effects as Russian 9-year-olds seem to be at a lower proficiency level.

Finally, past tense is used much more than 3SG-agreement with the exception of Chinese 9-year-olds who use it to a similar extent.

7.5.6 Individual results concerning main thematic verbs: what do they tell us about children's performance?

In this section, I focus on L2 learners' performance on main thematic verbs. Individual results can give us a better picture of the data, in this case of L2 children's performance providing information about the spread of the use of different forms. Data in this section are therefore presented for individual children.

Starting with Chinese 9-year-olds, Table 7.24 shows the total number of main thematic verb-tokens (range: 6-27) used per child in both stories sorted in declining order from largest to smallest number of verb-tokens. The table shows 1) considerable variability, 2) use of bare forms is very common, 3) as is the use of the periphrastic marking, and 4) no child performs at ceiling – i.e. inflecting all main verbs produced.

Table 7.24: Chinese 9-year-old children's performance data sorted by largest number of main verbs to smallest

N	Child ID	Correct	Incorrect: bare	Incorrect periphrastic	TOTAL main verbs
1	CH_GA_34	17	7	3	27
2	CH_GA_01	3	11	11	25
3	CH_GA_24	14	10	1	25
4	CH_GA_05	18	6	0	24
5	CH_GA_13	7	8	6	23
6	CH_GA_14	8	13	2	23
7	CH_GA_23	2	11	10	23
8	CH_GA_04	10	12	0	22
9	CH_GA_06	8	13	1	22
10	CH_GA_36	0	16	6	22
11	CH_GA_10	11	10	0	21
12	CH_GA_19	4	17	0	21
13	CH_GA_39	1	15	5	21
14	CH_GA_26	6	9	5	20
15	CH_GA_29	2	14	4	20
16	CH_GA_15	3	9	7	19
17	CH_GA_21	3	11	5	19
18	CH_GA_28	3	16	0	19
19	CH_GA_17	4	8	6	18
20	CH_GA_22	6	12	0	18
21	CH_GA_30	3	12	3	18
22	CH_GA_12	7	9	1	17

23	CH_GA_16	2	10	5	17
24	CH_GA_25	2	6	9	17
25	CH_GA_09	8	6	2	16
26	CH_GA_31	3	12	1	16
27	CH_GA_20	6	7	2	15
28	CH_GA_35	7	6	2	15
29	CH_GA_37	3	12	0	15
30	CH_GA_03	1	8	5	14
31	CH_GA_07	3	8	3	14
32	CH_GA_38	3	5	5	13
33	CH_GA_08	3	9	0	12
34	CH_GA_33	3	7	2	12
35	CH_GA_02	0	8	3	11
36	CH_GA_11	1	7	2	10
37	CH_GA_27	0	9	0	9
38	CH_GA_18	2	4	0	6
	TOTAL	**187**	**373**	**117**	**677**

Looking at older Chinese children's performance in Table 7.25, there are three interesting findings: 1) again there is great variability, 2) older children use many more verbs than younger ones (range between 16-31verbs), and 3) only one child reached 90% accuracy (in grey colour) and all produced at least some bare forms.

Table 7.25: Chinese 12-year-old children's performance data sorted by largest number of main verbs to smallest

N	Child ID	Correct: inflected	Incorrect: bare	Incorrect: periphrastic	TOTAL main verbs
1	CH_GB_04	16	13	2	31
2	CH_GB_32	15	12	1	28
3	CH_GB_01	17	10	0	27
4	CH_GB_22	3	19	4	26
5	CH_GB_12	15	9	1	25
6	CH_GB_27	11	13	0	24
7	CH_GB_02	13	9	1	23
8	CH_GB_07	2	19	2	23
9	CH_GB_16	21	2	0	23
10	CH_GB_30	1	20	2	23
11	CH_GB_06	15	6	1	22
12	CH_GB_13	13	6	3	22

13	CH_GB_20	16	6	0	22
14	CH_GB_28	12	8	2	22
15	CH_GB_03	5	12	3	21
16	CH_GB_14	13	4	3	20
17	CH_GB_08	11	9	0	20
18	CH_GB_09	12	8	0	20
19	CH_GB_21	8	12	0	20
20	CH_GB_23	6	14	0	20
21	CH_GB_10	9	9	1	19
22	CH_GB_17	11	7	1	19
23	CH_GB_25	3	16	0	19
24	CH_GB_29	15	4	0	19
25	CH_GB_05	13	5	0	18
26	CH_GB_11	2	16	0	18
27	CH_GB_33	9	9	0	18
28	CH_GB_15	1	16	0	17
29	CH_GB_19	7	8	2	17
30	CH_GB_26	7	10	0	17
31	CH_GB_34	13	2	2	17
32	CH_GB_18	9	7	0	16
33	CH_GB_24	3	12	1	16
34	CH_GB_31	7	7	2	16
	TOTAL	**334**	**339**	**34**	**707**

Turning to Russian 9-year-olds' performance as presented in Table 7.26, the use of main verbs ranges from 3 to 29 (where the largest value (29) seems to be an outlier -the next highest number of verbs used is 22). Those children's performance displays very low accuracy with no child performing at ceiling. Children's performance again is very diverse using all inflected, bare and periphrastic forms.

Table 7.26: Russian 9-year-old children's performance data sorted by largest number of main verbs to smallest

N	Child ID	Correct: inflected	Incorrect: bare	Incorrect: periphrastic	TOTAL main verbs
1	RU_GA_20	8	13	8	29
2	RU_GA_32	3	5	14	22
3	RU_GA_08	3	14	4	21
4	RU_GA_28	7	14	0	21
5	RU_GA_30	14	7	0	21

6	RU_GA_02	11	5	3	19
7	RU_GA_06	5	13	0	18
8	RU_GA_16	1	10	7	18
9	RU_GA_23	2	16	0	18
10	RU_GA_09	2	15	0	17
11	RU_GA_01	4	9	2	15
12	RU_GA_15	7	8	0	15
13	RU_GA_17	1	13	1	15
14	RU_GA_25	0	10	5	15
15	RU_GA_14	2	8	4	14
16	RU_GA_26	5	5	4	14
17	RU_GA_27	1	7	6	14
18	RU_GA_05	2	11	0	13
19	RU_GA_03	0	7	5	12
20	RU_GA_10	5	7	0	12
21	RU_GA_21	6	6	0	12
22	RU_GA_18	0	7	4	11
23	RU_GA_19	1	8	2	11
24	RU_GA_31	3	8	0	11
25	RU_GA_07	0	10	0	10
26	RU_GA_22	0	10	0	10
27	RU_GA_29	1	5	1	7
28	RU_GA_04	2	3	1	6
29	RU_GA_12	0	6	0	6
30	RU_GA_24	0	4	1	5
31	RU_GA_11	0	4	0	4
32	RU_GA_13	3	0	0	3
	TOTAL	**99**	**268**	**72**	**439**

Finally, Table 7.27 displays the performance on main verbs as produced by Russian 12-year-olds (ranging from 10 to 27). Specifically, there are 9 children (see grey coloured rows) who perform at ceiling, that is, they have more than 90% accuracy. (Only one Chinese 12-year-old also reached 90% accuracy while no other child in any group performed at ceiling.) Apart from those children, the rest still produce bare forms while there are 19/42 children who also use periphrastic marking.

Table 7.27: Russian 12-year-old children's performance data sorted by largest number of main verbs to smallest

N	Child ID	Correct: inflected	Incorrect: bare	Incorrect: periphrastic	TOTAL main verbs
1	RU_GB_19	18	8	1	27
2	RU_GB_03	22	3	0	25
3	RU_GB_08	25	0	0	25
4	RU_GB_11	8	8	9	25
5	RU_GB_27	4	16	5	25
6	RU_GB_07	21	2	0	23
7	RU_GB_25	17	6	0	23
8	RU_GB_05	14	8	0	22
9	RU_GB_15	20	1	1	22
10	RU_GB_22	6	15	1	22
11	RU_GB_34	6	13	3	22
12	RU_GB_02	12	7	1	20
13	RU_GB_37	1	20	0	21
14	RU_GB_10	3	15	2	20
15	RU_GB_30	14	6	0	20
16	RU_GB_18	18	0	1	19
17	RU_GB_36	17	2	0	19
18	RU_GB_31	19	0	0	19
19	RU_GB_26	15	3	0	18
20	RU_GB_39	4	14	0	18
21	RU_GB_12	16	2	0	18
22	RU_GB_13	5	7	6	18
23	RU_GB_16	11	6	1	18
24	RU_GB_23	4	14	0	18
25	RU_GB_41	4	13	1	18
26	RU_GB_21	11	5	1	17
27	RU_GB_29	16	1	0	17
28	RU_GB_33	17	0	0	17
29	RU_GB_14	12	4	0	16
30	RU_GB_17	2	9	5	16
31	RU_GB_38	6	9	1	16
32	RU_GB_01	9	6	0	15
33	RU_GB_04	12	3	0	15
34	RU_GB_06	13	2	0	15
35	RU_GB_42	6	9	0	15
36	RU_GB_20	4	8	2	14
37	RU_GB_28	12	2	0	14
38	RU_GB_32	13	1	0	14

39	RU_GB_35	12	1	0	13
40	RU_GB_09	3	4	5	12
41	RU_GB_40	4	3	4	11
42	RU_GB_24	0	10	0	10
	TOTAL	**456**	**266**	**50**	**772**

Summary of individual results:

Individual results confirm that the patterns we saw in the groups are reasonably spread across each group, so that individuals are not too different from the group average. We did not see bimodal distributions with children performing either at ceiling or very low, with the exception of the group of older Russian children in which many of them perform at ceiling and show they have achieved acquisition.

7.6 Summary of the main results of the chapter

- Inflectional morphology is a problematic domain for L2 children; low inflection rates on main thematic verbs were found for all groups while copula and auxiliary *be* showed that even those morphemes are not totally unproblematic.
- Developmental effects: Older children are generally more accurate than younger ones and their performance is less diverse.
- L1 effects: Group results as well as individual results revealed differences between L1s with Russian 12-year-olds outperforming Chinese 12-year-olds in use of inflection on main thematic verbs in narratives; while there were 9 Russians who performed at ceiling (≥90%) inflecting the majority of the main thematic verbs they used, there was only one Chinese 12-year-old reaching 90% accuracy.
- Morphemes: Children mark for past much more than they do for 3SG-agreement.
- Morphemes: Learners do better with copula *be* than with auxiliaries, and in both better than with main thematic verb inflection.
- Morphemes: 3SG contexts seem to be easier for learners than 3PL contexts which appeared quite challenging.
- Individual results showed wide accuracy ranges with respect to finiteness marking for all groups. CH_9: 0-18, CH_12: 1-21, RU_9: 1-14, RU_12: 0-25. However, it is interesting that only some of the Russian 12-year-old group performed at ceiling in narratives correctly inflecting verbs 90% of times.

In the next chapter, I focus on the periphrastic structure '*be + verb(x)*' attested in both TEGI tasks and narrative production. In Chapter 9, I present statistical analyses and discuss results altogether in Chapter 10.

CHAPTER 8

THE PERIPHRASTIC STRUCTURE

8.0 The pattern '*is + verb(x)*'

In this chapter, I focus on the periphrastic structure '*is + verb(x)'* which was attested in both TEGI tasks and narrative production. It is a pattern produced by immersed L2 children (e.g. Ionin & Wexler, 2002; Paradis et al., 2008) as well as by instructed learners; it was found in written narratives by Cantonese learners of English in Hong Kong classrooms (Yang & Huang, 2004) as well as in oral narratives by Spanish children classroom learners of English (e.g. Garcia-Mayo et al., 2005). What makes this pattern interesting is that it is not part of the input so, it cannot be the result of rote learning. It could be the result of mis-chunking of the input or some incomplete form, e.g. children actually attempting a present progressive form 'is catching' instead of 'is catch'. But if we can show that this is not the case, then, it can potentially shed light to the internal processes involved in the acquisition of finiteness features.

In this chapter, I consider the empirical properties of this pattern (sections 8.1-8.4) looking at its frequency and forms, agreement, tense and the lexical aspect used in periphrastic patterns and discuss it in detail in Chapter 10.

8.1 Frequency and forms

Table 8.1 shows the use of the periphrastic construction in TEGI. The frequency of the periphrastic marking instances out of all verbs in TEGI tasks is given in the grey coloured columns.

Chinese 9-year-olds tend to use '*be + bare verb*' (e.g. '*A dentist is fix teeth'*) in both TEGI 3SG and past, whereas their Russian counterparts also use the '*be + bare verb*' in the TEGI 3SG, but in TEGI past, they mostly use '*be + inflected verb*' (e.g. '*He is brushed his hair'* or '*She is climbed*.'*)*.

Chinese 12-year-olds also use the '*be + bare verb*' structure in both TEGI tasks but the raw occurrences are few. As for Russian 12-year-olds, they also use the '*be + bare verb*' in

TEGI 3SG but mainly '*be + inflected verb*' (as in '*He was painted'*) in TEGI past. It seems that '*be + inflected verb*' is the most frequent pattern from Russians in TEGI past.

Table 8.1: Raw instances and percentages of periphrastic forms with be in both tasks for all groups across the two L1 cohorts

Forms in TEGI tasks		TEGI 3SG		TEGI Past		
		'be + bare V'	***'be + bare V'/ all verbs***	***'be + bare V***	***'be + inflected V'***	***'be + V(x)'/all verbs***
Chinese	9yo.	77	77/370 21%	56/85 **66%**	29/85 34%	85/576 15%
	12yo.	16	16/340 5%	15/31 **48%**	16/31 **52%**	31/612 5%
Russians	9yo.	65	65/310 21%	51/141 36%	90/141 **64%**	141/522 27%
	12yo.	36	36/420 8.5%	18/77 23%	59/77 **77%**	77/756 10%

Turning to narratives, Table 8.2 presents the three categories of forms of this structure (i.e. '*be + bare verb, be + verb inflected for 3SG,* and *be + verb inflected for past'*) and the percentages of their production by each group in narrative production. The frequency of the periphrastic marking instances out of all main thematic verbs are given in percentages in the grey coloured column. Examples of each category respectively are: *'dog is run', 'The boy is runs', 'The boy is used the …'.* Recall that in case of narratives, I did not score forms in obligatory contexts but overt finiteness marking (see Chapter 7, section 7.1); hence, in this task there is no distinction between present and past tense contexts.

Chinese 9-year-olds almost exclusively use '*be + bare verb*' (95%). Russian 9-year-olds also use this form very extensively (85%) but they also produce a small number of instances of '*be + verb inflected for past*' (15%).

The pattern persists in the productions of Chinese 12-year-olds amounting to 70.5% of periphrastic uses. Other uses emerge as well, mainly '*be + inflected for past verb*' (23.5%). The '*be + bare verb*' also persists with Russian 12-year-olds as the main pattern of periphrastic verb marking but its use decreases to 61.5% while the use of '*be + inflected for past verb*' doubles reaching 32.7%.

In all cases the '*be + inflected for 3SG verb*' is very rare. The differences between the groups are not very pronounced as all groups mainly use '*be + stem verb*'. However, Russians use *'is + inflected for past verb'* more frequently if we compare the younger Chinese and Russian groups as well as the older groups between them. Older children also use '*be + inflected for past verb*' more than younger ones.

Table 8.2: The forms the periphrastic structure appears in

Forms in narratives	**be + bare verb**	**be + inflected for past verb**	**be + inflected for 3SG verb**	**Total instances of periphrastic marking**	**Periphrastic marking / main thematic verbs**
CH_9	111/117 **94.9%**	2/117 1.7%	4/117 3.4%	117 100%	117/677 17.3%
CH_12	24/34 **70.5%**	8/34 23.5%	2/34 6%	34 100%	34/707 5%
RU_9	62/73 **85%**	11/73 15%	0/73 0%	73 100%	73/439 16.5%
RU_12	32/52 **61.5%**	17/52 32.7%	3/52 5.8%	52 100%	52/772 6.5%

Thus, in both production tasks, we see that the form '*be + bare verb'* prevails in all groups, with the exception of Russians in past; they produce '*be + inflected for past verb'* more in past contexts. A difference between the two tasks is that Russian children use the '*be + inflected for past verb'* much less in narratives than in TEGI perhaps because there are no obligatory past contexts. Finally, the frequency of the periphrastic marking instances in both TEGI tasks and narratives shows that it is a robust pattern especially in younger children's production.

8.2 The pattern ‘*be + verb(x)*’: Agreement

The next question is whether ‘*be*’ in the periphrastic structure agrees with its preceding subject. In the TEGI task, prompts consisted only of 3SG contexts so we could only look for agreement errors on forms of auxiliary ‘*be*’ in this particular context. Table 8.3 shows the number of instances of production of incorrect use of ‘*are*’ or ‘*were*’. In both L1 groups, there were very few children who made such errors.

Table 8.3: Raw number of instances of production of agreement error in ‘be’ forms of the structure ‘be + verb’

Agreement		**TEGI 3SG**	
		incorrect: plural	produced by
CH	9	5	1 child
	12	1	1 child
RU	9	5	4 children
	12	7	1 child

Turning to narratives, Table 8.4 shows the raw numbers and percentages of (incorrect) agreement between the subject and the form ‘*be*’ of the periphrastic structure. Results are provided by number on subject; that is, 3SG when the noun preceding the verb is 3SG and 3PL for the plural nouns.

It is quite clear that agreement marking is unproblematic in the vast majority of the cases across all groups. There were very few cases of incorrect agreement such as ‘*A mother goat and her second daughter was drink water’ or ‘his ball are fall to lake’*. Children did not use many plural subjects in general. It is, therefore, not possible to conclude reliably that agreement is established only on the basis of singular subjects. Recall that when considering children’s performance on copula *be* when there were plural subjects, accuracy dropped considerably (see Tables 7.9 and 7.11 in Section 7.5.1) and we saw utterances such as ‘**the sausages is in the bag*’ or ‘*... and the goats is happy’*.

Table 8.4: Raw numbers and percentages of correct and incorrect agreement between subject and the verb 'be' in periphrastic structures in narratives

Agreement	Subjects	Correct agreement of aux *be*	Incorrect agreement of aux *be*	Total number of instances of *'be + verb'*
CH_9	3SG	113/117 **96.5%**	4/117 **3.5%**	117 100%
	3PL	-	-	-
CH_12	3SG	32/34 **94%**	2/34 **6%**	34 100%
	3PL	-	-	-
RU_9	3SG	70/72 **97%**	2/72 3%	72 100%
	3PL	1/1 **100%**	0/1 **0%**	1 100%
RU_12	3SG	48/50 **96%**	48/50 **4%**	50 100%
	3PL	0/2 **0%**	2/2 **100%**	2 100%

8.3 The pattern 'is + verb(-x)': tense and number features

I now turn to explore whether '*be*' was marked for tense. Starting again with the TEGI tasks, I only tested for tense features in case of TEGI past.[26] Table 8.5 presents the raw numbers as well as percentages of forms of '*be*' of the periphrastic structure appearing as '*is'* or '*was'*. Recall that there were no 3PL contexts in TEGI tasks. As can be seen younger children produce almost exclusively *'is'* while older ones use also '*was*' to a larger extent even if '*is*' still prevails. It seems then that *'is'* is used for both present and past contexts by younger children.

[26] When scoring TEGI 3SG, I saw that use of past on auxiliary '*be'* of the periphrastic structure was rare to non-existent. Hence, I decided there was no reason to look for past tense features in this task.

Table 8.5: Raw numbers and percentages of forms of ‘be’ with respect to tense marking of the periphrastic structure in TEGI past task

Tense	TEGI Past		
	‘Is’	**‘Was’**	**Total instances of ‘be + verb(x)’**
Examples:	*A dancer is dance.*	*He was brush(ed) his hair.*	
CH_9	110 **83%**	22 **17%**	132
CH_12	42 **62%**	26 **38%**	68
RU_9	218 **96%**	9 **4%**	227
RU_12	69 **69%**	31 **31%**	100

Turning to narrative production, Table 8.6 presents the raw numbers and percentages of use of the different forms of ‘*be’* as produced by each group. As there were just 3 plural subjects preceding this structure, I will consider both *‘is’/’are’* together and *‘was’/’were’* together.

Chinese and Russian 9-year-olds mostly use ‘*is’* with Chinese 9-year-olds using the past form ‘*was’* slightly more frequently than the Russian 9-year-olds. Note however that the raw numbers were very few.

Chinese 12-year-olds use ‘*is’/’are’* in 70.5% percent of the cases while they also use ‘*was’* quite frequently (29.5%). Russian 12-year-olds mostly use ‘*is’* (77%) while they also produced some instances of *‘was’/’were’* (23%).

Table 8.6: Raw numbers and percentages of different persons/forms of '*be'* as produced by each age and L1 group

Tense	number of instances of **'is'** / **'are'**	number of instances of **'was'/ 'were'**	**Total**
CH_9	109/117 **93%**	8/117 **7%**	117 **100%**
CH_12	24/34 **70.5%**	10/34 **29.5%**	34 **100%**
RU_9	72/73 **98.6%**	1/73 **1.4%**	73 **100%**
RU_12	40/52 **77%**	12/52 **23%**	52 **100%**

Based on these results, we conclude that there is very little past marking in younger children who mostly use *'is'* and not *'was'*. Older children seem to use the forms *'was'(/'were')* to a larger extent, however, we should be interpreting these results with caution because the number of periphrastic marking instances for older groups is quite small. Finally, it seems that as tense marking increases as it happens for older children, the use of periphrastic structure is being reduced.

I now turn to a further empirical question, the lexical aspect of verbs marked by the periphrastic construction.

8.4 The lexical aspect of the periphrastic structure

The data from TEGI showed that the periphrastic structure appeared with a variety of verbs (telic and atelic), but as the sample of verbs was relatively small, we could not investigate any potential correlations between lexical aspect and the periphrastic structure. I thus consider this issue in the narrative production.

To consider the lexical aspect of the periphrastic structure, we need to compare it with the lexical aspect of all inflected verbs, that is, main thematic verbs following 3SG subjects (N=2,322), the periphrastic structure in both 3SG and 3PL contexts (N=276), as well as progressive tense marking in both singular and plural contexts (N=485), thus, in total, 3,083 items.

Let us first look at the aspectual classes of all predicates as represented by our verb-tokens used by the children. Table 8.7 shows that the most frequent type is achievements, accounting for just over 40% of uses (1,299/3,039). Most frequent achievement verbs in our data are: *see, catch, fall, notice, drop,* etc. Achievements are followed by accomplishments (763/3,039) such as '*fly away, go back, climb the tree, eat a sausage*' accounting for a quarter of cases. Statives correspond to 17% of uses (521/3,039) with '*want'* the most common instantiation. Finally, activities were the smallest category reaching 15% (456/3,039), closely behind statives. Predicates such as '*fly, swim, drink, look, watch*' are some common examples of activity verbs. A total of 44 cases out of the total 3,083 were not coded for aspect (NAs)[27] (cf. Section 7.1: Verb annotation for lexical aspect) and were excluded from analysis: 38 main thematic verbs and 6 verbs in the periphrastic structure.

Table 8.7: Children's total production of verbs in the narrative task and their distribution over aspectual classes

Stative	**Activity**	**Accomplishment**	**Achievement**	**Total**	**NAs**
521	456	763	1,299	3,039	44

Let us now consider lexical aspect and inflectional marking. Table 8.8 shows the number of instantiations of each suffix (3SG *-s*, past *-ed* (or irregular), progressive *-ing*) with verb-tokens of each lexical aspectual class also including (incorrect) bare forms (null ∅) and the periphrastic structure. The green color shows percentages of each morpheme/structure for each aspectual class (i.e. within column) and the blue color shows percentages of lexical class for each morpheme (i.e. across the row).

[27] NAs stand for 'Not Available', that is, missing values in R coding. Recall that some verb-tokens were not determined for lexical aspect as explained in Chapter 7, Section 7.1.

Table 8.8: Raw numbers and percentages of morphemes, bare forms and periphrastic structure instances classified by aspectual class

	Stative	**Activity**	**Accomplishment**	**Achievement**	**Total**
3SG	121 42.5% 23%	28 10% 6%	37 13% 5%	99 34.5% 7.5%	285
past	104 13.5% 20%	75 9.5% 16.5%	147 19% 19%	452 58% 35%	778
progressive	10 2% 2%	181 38% 39.5%	196 41% 25.5%	93 19% 7%	480
bare	270 22% 52%	132 11% 29%	281 23% 37%	543 44% 42%	1,226
periphrastic	16 6% 3%	40 14.8% 9%	102 37.7% 13.5%	112 41.5% 8.5%	270
Total	521	456	763	1,299	3,039

We see that bare forms are the most common form for all verb types, except activities where progressive is the most frequent form. Turning to morphemes, 3SG-agreement is most frequent with statives and achievements while past morphology is most frequent with achievements. Progressive is high with activities and accomplishments. Periphrastic marking is used mainly with achievements and accomplishments. Figure 8.1 offers a schematic presentation of these results.

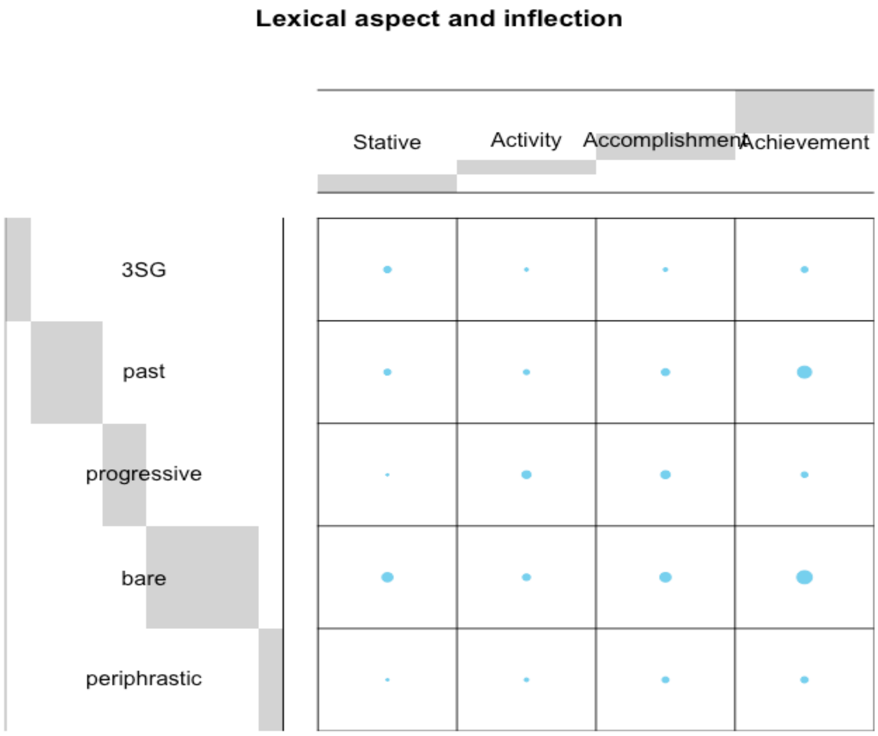


Figure 8.1: Illustration of associations between morphology and lexical aspectual classes. (The size of the dots represents the magnitude of the association. The grey coloured areas show the frequency of the morphology or the aspectual class in the data. In this case, for example, bare is the most common type of verb followed by past, etc.[28]*)*

This picture as described above and as shown in Figure 8.1 demonstrates some clear links between morphology and the lexical semantics. Let us now consider if these correlations between lexical aspect and verbal morphology are similar across Russians and Chinese and across ages.

Table 8.9 shows that Chinese 9-year-olds produced in total 782 verb-tokens (5 verb-tokens were not coded for aspect and were excluded) being ordered in the same way as that of the general pattern we saw above (see Table 8.7).

Table 8.9: Chinese 9-year-olds' production of verbs in the narratives task and the number of their classification into aspectual class

Stative	Activity	Accomplishment	Achievement	Total	NAs
173	120	186	298	777	5

[28] This holds for all other figures of this chapter which are of the same type.

Table 8.10 further shows the use of each particular suffix (3SG, past, progressive, bare) or periphrasis across aspectual verb classes.

Table 8.10: Suffixes, bare forms and periphrastic occurrences by lexical classes of verbs in raw numbers and percentages

	Stative	Activity	Accomplishment	Achievement	TOTAL
3SG	43 47% 25%	11 12% 9%	9 10% 5%	28 31% 9%	91
past	10 10.5% 6%	13 13.5% 11%	21 22% 11%	51 54% 17%	95
progressive	1 1% 0%	35 33.3% 29%	38 36.2% 20.5%	31 29.5% 10.5%	105
bare	111 30% 64%	42 11.3% 35%	68 18.2% 36.5%	151 40.5% 51%	372
periphrastic structure	8 7% 5%	19 16.5% 16%	50 44% 27%	37 32.5% 12.5%	114
TOTAL	173	120	186	298	777

By and large, data follow the general pattern apart from two cases: 1) activities are mostly bare rather than inflected for progressive, 2) the periphrastic marking mainly occurs with accomplishments (44%) followed by achievements (32.5%) following the reverse pattern from the general pattern where achievements were slightly more common. Figure 8.2 illustrates these results schematically.

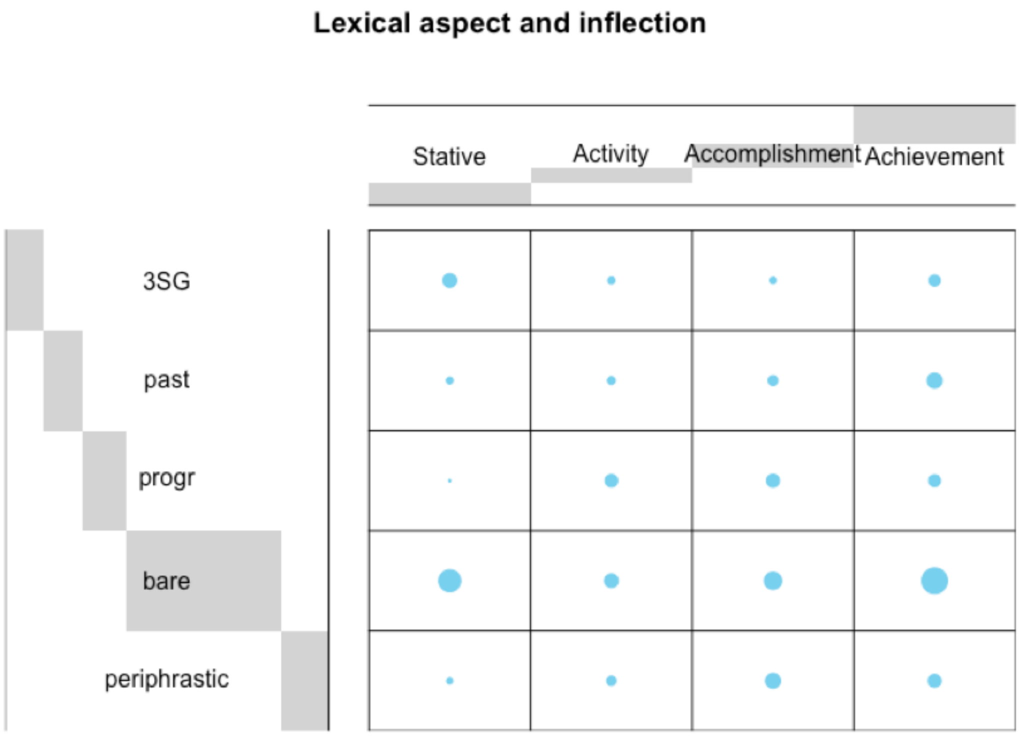


Figure 8.2: Graphical display of the relationships between morphology and aspectual class of the verb tokens produced by Chinese 9-year-olds

Table 8.11 shows Chinese 12-year-olds' production of verbs and their classification into aspectual classes. 12-year-old children produced in total 812 verb-tokens (from which 12 were not coded for aspect and were excluded), these being classified in aspectual categories in the exact same order as that of the general pattern (Table 8.7).

Table 8.11: Chinese 12-year-olds' production of verbs in the narratives task and the number of their classification into aspectual classes

Stative	**Activity**	**Accomplishment**	**Achievement**	**Total**	**NAs**
136	71	202	391	800	12

Table 8.12 provides the distribution of inflected, bare and periphrastic forms into aspectual classes.

Table 8.12: Number of inflected and bare forms for verbs in each aspectual class as produced by Chinese 12-year-olds

	Stative	**Activity**	**Accomplishment**	**Achievement**	**Total**
3SG	34 28.5% 25%	13 11% 18%	15 12.5% 7.5%	57 48% 14.5%	119
past	25 11.8% 18.4%	4 1.9% 6%	42 19.8% 21%	141 66.5% 36%	212
progressive	3 3% 2.2%	32 31% 45%	37 36% 18%	31 30% 8%	103
bare	74 22.2% 54.4%	20 6% 28%	91 27.3% 45%	148 44.5% 38%	333
periphrastic structure	0 0% 0%	2 6% 3%	17 51.5% 8.5%	14 42.5% 3.5%	33
TOTAL	136	71	202	391	800

Again, data are very similar to the general pattern. The main major differences are: 1) 3SG-agreement is more frequent with achievements, followed by statives in contrast to the general pattern where 3SG-agreement occurs mainly with statives. 2) Progressive aspect is found to be produced to a similar extent with all aspectual classes except statives (very few instances), a pattern that Chinese 9-year-olds also showed but which is different from the general pattern in that achievements are less marked for progressive. 3) Finally, periphrasis is used mainly with accomplishments (51.5%) and achievements (42.5%) –which is slightly dissimilar to the general pattern (Acc.: 37.7% vs Ach.:41.5%). Figure 8.3 presents the results schematically.

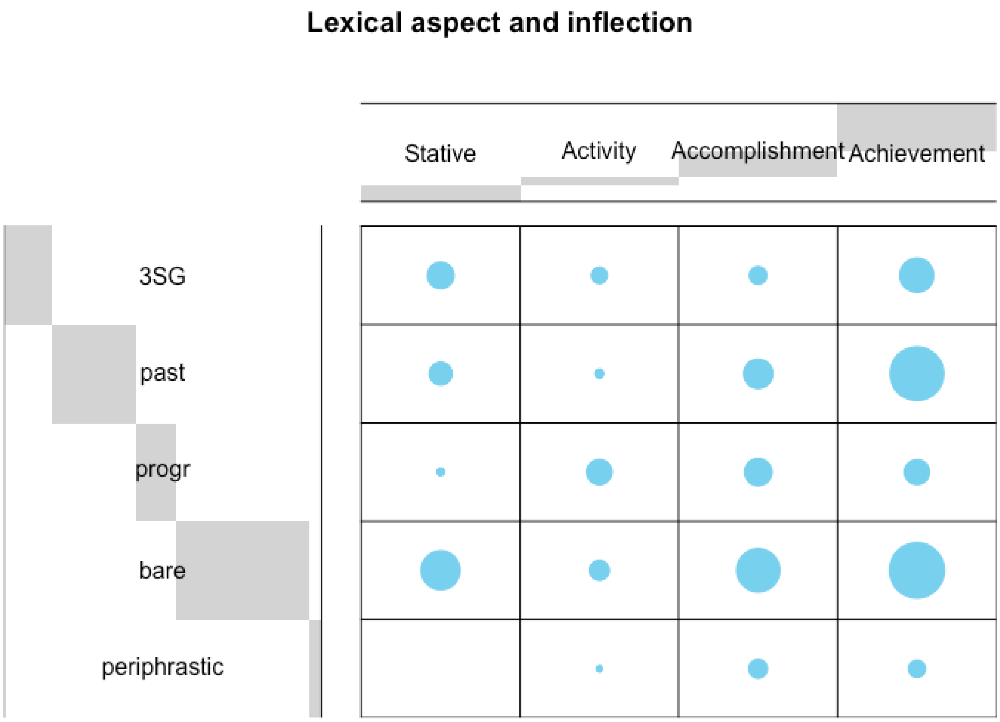


Figure 8.3: Graphical display of the relationships between morphology and aspectual class of the verb tokens produced by Chinese 12-year-olds

In sum, Chinese children show similar correlations with the overall pattern except in three cases; for younger children bare forms are the most frequent choice across all classes, including activities. Older children deviate from the general pattern in using most 3SG-agreement with achievements rather than statives. Also, Chinese groups use periphrasis more frequently (10% more) with accomplishments than achievements (slightly) contrasting the general pattern which shows a smaller difference (4%) between the two.

Russian 9-year-olds produced fewer than all other groups verb-tokens summing up to 572 instances. From those, 6 instances were not coded for aspect. The frequency of verb classes diverges slightly from the general pattern in that younger Russians use more activities than statives as shown in Table 8.13 (compared to Table 8.7).

Table 8.13: Russian 9-year-olds' production of verbs in the narratives task and the number of their classification into aspectual class

Stative	Activity	Accomplishment	Achievement	Total	NAs
66	109	183	208	566	6

Table 8.14 presents the raw numbers as well as percentages of the verbs of each aspectual class being marked or not.

Table 8.14: Number of inflected and bare forms for verbs in each aspectual class as produced by Russian 9-year-olds

	Stative	**Activity**	**Accomplishment**	**Achievement**	TOTAL
3SG	19 63% 29%	1 3% 1%	5 17% 2.7%	5 17% 2%	30
past	5 7.5% 7.5%	15 22.4% 13.7%	16 23.8% 8.7%	31 46.3% 15%	67
progressive	0 0% 0%	50 38% 45.8%	68 52% 37.1%	13 10% 6%	131
bare	37 14% 56%	31 11.7% 28.5%	73 27.5% 40%	124 46.8% 60%	265
periphrastic structure	5 7% 7.5%	12 16% 11%	21 29% 11.5%	35 48% 17%	73
TOTAL	66	109	183	208	566

As can be seen there are no major differences from the general pattern. It should be noted though that the periphrasis is used considerably more with achievements (48%) than with accomplishments (29%) showing a much bigger difference between them compared to the difference attested in the general pattern (41.5% vs. 37.5% respectively). Figure 8.4 displays the results schematically.

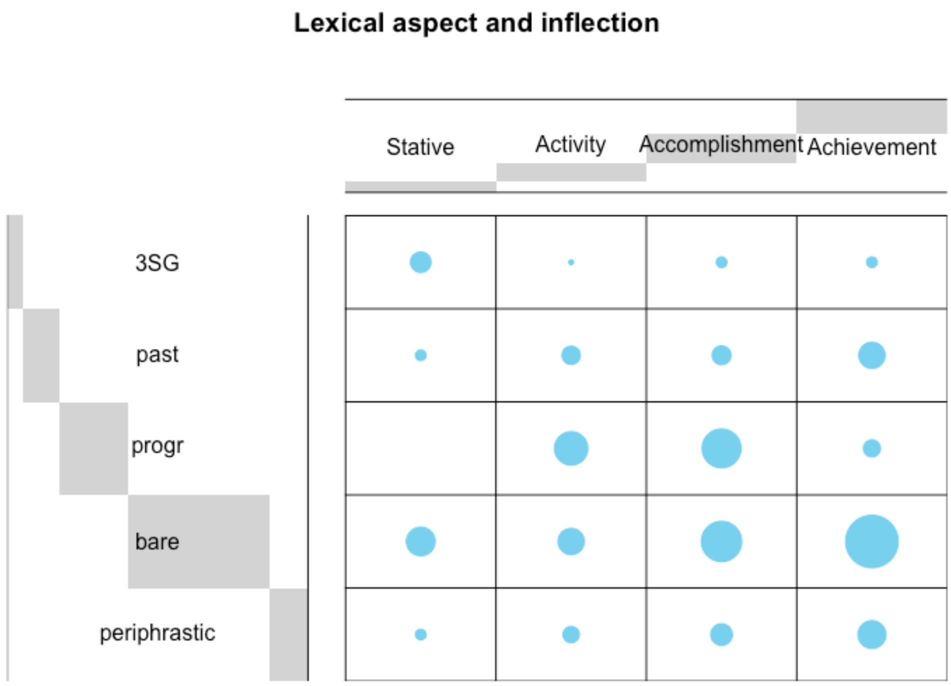


Figure 8.4: Graphical display of the relationships between morphology and aspectual class of the verb tokens produced by Russian 9-year-olds

Finally, with respect to the Russian 12-year-olds out of the 917 verb-tokens in their narratives, 21 were not coded for aspect and were excluded. As can be seen in Table 8.15, older Russians also use more activities than statives, like the younger Russians, a divergence from the general pattern in Table 8.7.

Table 8.15: Russian 12-year-olds' production of verbs in the narratives task and the number of their classification into aspectual class

Stative	**Activity**	**Accomplishment**	**Achievement**	**Total**	**NAs**
146	156	192	402	896	21

Table 8.16 presents inflected, periphrastic and bare forms within each aspectual class.

Table 8.16: Number of inflected and bare forms for verbs in each aspectual class as produced by Russian 12-year-olds

	Stative	**Activity**	**Accomplishment**	**Achievement**	TOTAL
3SG	25 57% 17%	3 7% 2%	8 18% 4.2%	8 18% 2%	44
past	64 16% 44%	43 10.5% 27.5%	68 16.7% 35.5%	230 56.8% 57.2%	405
progressive	6 4% 4%	64 45% 41%	53 38% 27.5%	18 13% 4.5%	141
bare	48 19% 32.5%	39 15% 25%	49 19% 25.5%	120 47% 29.8%	256
periphrastic structure	3 6% 2%	7 14% 4.5%	14 28% 7.3%	26 52% 6.5%	50
TOTAL	146	156	192	402	896

In contrast to the general pattern (Table 8.8), for older Russian children bare forms are the minority of forms across all classes. The periphrastic marking is mostly used with achievements (52%) followed by accomplishments (28%) – a pattern similar to Russian 9-year-olds' and the general pattern but in contrast to the Chinese groups which show the reverse. Figure 8.5 offers a schematic presentation of the results.

Lexical aspect and inflection

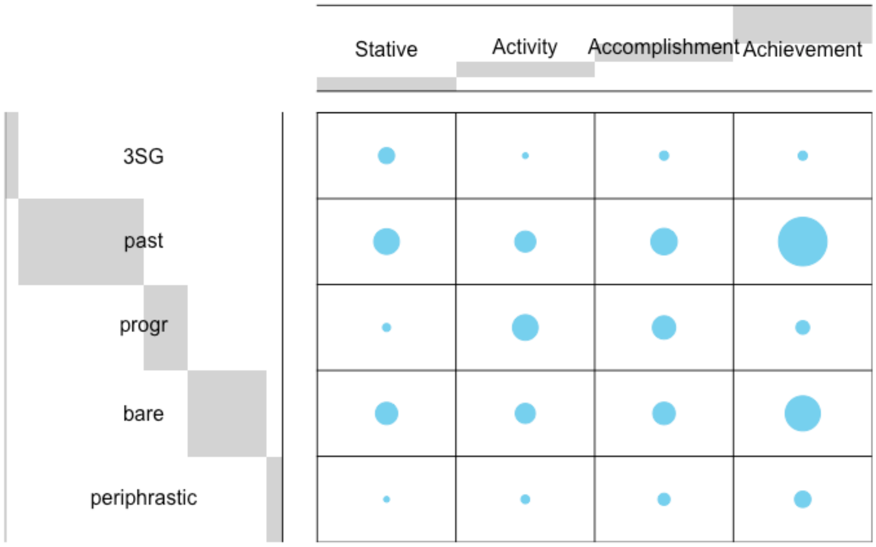


Figure 8.5 : Graphical display of the relationships between morphology and aspectual class of the verb tokens produced by Russian 12-year-olds

8.4.1 The lexical aspect of the periphrastic structure. Summary

Based on the above presentation, we saw that there are associations between morphology and lexical aspectual classes. Bare forms were the most common types of verbs with all aspectual classes except for activities which were mainly marked for progressive. Past is highly associated with achievements and 3SG-agreement with statives and to a lesser extent with achievements. Progressive is used mainly with accomplishments and activities. Finally, periphrastic marking is occurring mostly with accomplishments and achievements.

8.5 Summary of results of the chapter

This chapter aimed to explore the properties of the periphrastic structure '*is + verb(x)*' based on production data by Chinese and Russian EFL learners. Combining the empirical observations from both tasks, the periphrastic structure is characterised by:

1. use of inflected '*be'* forms such as '*is/was'* (and less '*are/were'*) with both bare and inflected verb-forms (*e.g. 'is play', 'is jumped'*)
2. most frequent/dominant use as '*is'* showing agreement in 3SG contexts; however, it is almost never used with plural subjects

3. lack of evidence for tense features; especially 9-year-olds use almost always '*is'* and not '*was'* in past tense contexts*;* older children use '*was*' more but the numbers are very small
4. use of the periphrastic structure mainly with accomplishments and achievements.

I discuss these results in Chapter 10 (Discussion) in connection to the relevant hypotheses. In the next chapter, I present statistical analyses for both TEGI tasks and narratives.

CHAPTER 9

STATISTICAL MODELLING FOR FINITENESS MARKING IN TEGI TASKS AND NARRATIVE PRODUCTION

9.0 Introduction on statistical modelling

In this chapter, I present the statistical analyses of the results with respect to finiteness marking in the two studies, TEGI tasks and narrative production. The main question is to identify the predictors of accuracy/overt finiteness marking. I used mixed-effects models which I introduce in some detail below and then I present the results for each analysis. Before so, I will go through the variables included in the models.

9.1 Variables included in the analyses

Table 9.1 displays the dependent variable (i.e. accuracy) as well as all the predictor/independent variables, their coding, and the levels they consist of. Not all of them were included in each analysis; I discuss the structure used for each one depending on its aims.

Table 9.1: The variables, their coding, and the levels they consisted of

Variable	Coded as	Levels
accuracy	acc	0 or 1
L1	nation	CH or RU
age	age	9 or 12
(Renfrew) vocabulary score	(z)voc	continuous
academic hours of attendance in EF	(z)hours	continuous
inflection	tense	3SG or past
regularity	reg	regular or irregular
narrative length	(z)nar_length	continuous
number	number	3SG or 3PL
verb type	verbtype	cop be, aux be, main thematic verb

Starting with the dependent variable, this was always accuracy, coded as 0 for incorrect answers and 1 for correct answers. To give an example, in TEGI-3SG task, if a child said 'He plays...', this item would be coded as 1, otherwise as 0.

Turning to independent variables, age is a categorical variable in this study having two levels 9 or 12. Similarly, L1 has two levels Chinese or Russian. The Renfrew vocabulary score and narrative length were continuous measures which were considered as a proxy for proficiency. The academic hours of attendance in EF were taken as a measure of exposure which was included in order to control for input/exposure effects as I observed a discrepancy in mean academic hours between 9-year-olds in the two countries (see Table 5.8, Section 5.2.5). Inflection coded (or labelled) as tense was added in analysis considering whether there is an asymmetry in children's performance between the two morphemes, 3SG-agreement and past tense; in other words, the inflection variable has two levels 3SG or past. Regularity concerned only TEGI past and concerned the regularity status of the verbs: regular or irregular. Number is relevant in narratives and specifically, copula *be* and auxiliary *be*. (Plural number in main verbs is excluded from the analyses.) Finally, verb type also applied only in narratives where children produced apart from main thematic verbs also the copula and the auxiliary *be*.

Two more variables were not included in the models to avoid collinearity. The first variable is age of onset as distinct from age of testing. Since I chose children which had 5 years of EF instruction, age of onset and age of testing in effect capture the same aspect.[29] The second variable is the CEFR level of the class children attended at time of testing, because it would correlate highly with children's age as most 9-year-olds attended A2 level classes and the vast majority of the 12-year-olds attended B1 level classes.

9.2 Linear mixed-effects models

Mixed-effects logistic regression (MELR) analyses were conducted in the R software (R Core Team, 2014) using the glmer (generalized linear mixed effects regression) function from the package lme4 (Bates, Maechler, Bolker, & Walker, 2015). In essence, mixed effects models can answer the same questions as regression or ANOVA, however, I opted for mixed-effects models over other analyses because they allow to account for variance in the data arising from

---

[29] There was small variability though within this variable as some parents reported an earlier age of onset of their children's onset of learning English (13 Chinese 12-year-olds, 6 Russians 12-year-olds) either in EF or outside EF. The small variability in the age of onset variable led me to decide not to include this variable in the analyses.

random population and/or item sampling. In other words, they allow us to consider control predictor variables in a single analysis whereas traditional methods would require additional analyses (Cunnings, 2012; Baayen et al. 2008). As the dependent variable accuracy, in these analyses is always coded as binary (0 or 1), 0 for incorrect and 1 for correct use of the relevant morpheme (e.g. 3SG -*s* or Past -*ed*), binomial logistic regression[30] analyses were used.

Before running the statistical analyses, values of continuous predictor variables were z-score transformed (i.e. having mean of 0 and standard deviation of 1) 'to eliminate potential scale bias' (Schütze & Sprouse, 2013:43). This is a common step in data pre-processing of scale-type data. Z-score transformation makes variables previously measured in different scales more comparable since they are converted into standard units and results are more interpretable.

The methodology of the analysis consisted of the following steps: I first applied a full fixed effects structure which accounted for as many main effects and interactions as possible. I applied to this structure, various random structures in order to determine the best random structure for the full model. I carried out model comparisons through the anova function in order to find out which structure provided the lowest AIC (i.e. Akaike Information Criterion) which is a measure of the relative quality of the statistical models. The smaller the AIC the more variance our model can explain.

Subsequently, I conducted a backward stepwise model comparison using Maximum Likelihood (ML) and dropping effects of the fixed structure that were not significant predictors of variance in the data and were not involved in any significant interaction. I started checking the three-way interactions, proceeding to two-way interactions and finally, considered the main effects. Only one interaction or main effect was dropped each time. A new full model was then created and new comparisons were carried out. In order to determine which model was best-fitting, I used the likelihood ratio test with the anova function (Cunnings, 2012). In all analyses, the level of significance was taken to be $p<.05$.

In case of significant three-way interactions, data were divided into two subsets according to L1 and adjusted LMMs (i.e. Linear Mixed Models) were fitted for the two subsets. To explore two-way interactions, I ran post-hoc tests using the testInteractions function of the phia R package (De Rosario-Martínez, 2013).

---

[30] When the outcome of the dependent variable is binary/dichotomous, then binomial logistic regression should be used.

Finally, logistic regression assumes that there is no collinearity between the variables. Most of the variables I included in my models, i.e. age, L1, vocabulary size, and academic hours of attendance in EF are different measures of separate constructs exerting different influences on morpheme accuracy, such that the assumption is typically met. When this is not the case, I report correlational analyses.

9.3 Statistical analyses for TEGI tasks

Starting with the TEGI tasks, the main question is what predicts accuracy in 3SG and past tense (regular and irregular); the age of the learners, their L1, their vocabulary score, or just the academic hours they attended in EF. Three separate analyses were conducted for accuracy in 3SG-agreement, past regular and past irregular tense and the same fixed structure was applied in all three cases. The two separate analyses for past tense, regular and irregular verbs can clarify if the same factors predict accuracy for both types of verbs as theory has discussed the possibility that their learning tap into different types of learning.

Specifically, the full fixed structure that was applied -with accuracy in 3SG/regular past /irregular past as the dependent variable included age, L1, vocabulary score and the academic hours in EF, structured as two potential three-way interactions (age x L1 x vocabulary + age x L1 x hours). The structure further contains all lower level two-way-interactions as well as the main effects. This structure tells if vocabulary differs across ages and L1s and whether it can predict accuracy on 3SG-agreement or past tense for one, some or all groups. Academic hours in EF were added in a separate construct to investigate whether if differing across age groups and L1s could explain accuracy.

The best random structure in all three models was the simplest one including a random intercept for participants and a random intercept for items.

Three Chinese children, one 9-year-old and two 12-year-olds (CH_GA_35, CH_GB_23, CH_GB_25) were excluded from the statistical analyses as there were missing data concerning their academic hours of attendance at EF.

These analyses were followed by two more. The first one aimed to answer whether the regularity of the verbs could predict accuracy in past. In this analysis, the fixed structure included a three-way interaction between L1, age, and regularity (L1 x age x regularity). This structure can clarify whether regularity predicts accuracy in past tense and whether L1 and age groups show different patterns to this respect. The best random structure was a random intercept for participant and item.

The final analysis of TEGI tasks carried out aimed to explore whether there is a significant difference in children's accuracy on 3SG-agreement versus past tense. To address the question whether 3SG-agreement is harder to acquire than tense I again carried out a mixed effects logistic regression to compare children's performance on 3SG and regular past tense. I excluded irregular past from this analysis for the following reason: while regular past requires the application of a grammatical rule, the irregular forms can be learnt using different means as they are suppletive. Thus, the 3SG *-s* morpheme is compared with the *-ed* past morpheme. To establish whether 3SG-agreement is harder than past tense, this needs to be the case across ages and L1s. Considering again accuracy as the dependent variable, I included the variables L1, age, and inflection -coded as tense- as the predictor variables. Tense was a categorical variable with two levels: 3SG-agreement and past tense. The maximal model included a three-way interaction between L1, age, and tense (i.e. L1 x age x tense). The best random structure was one that specified a random slope for age by participant and a random intercept for item (age|part) + (1|item).

9.3.1 Binomial logistic regression for TEGI 3SG

Figure 9.1 presents the mean accuracy results on 3SG-agreement for each age and L1 group.

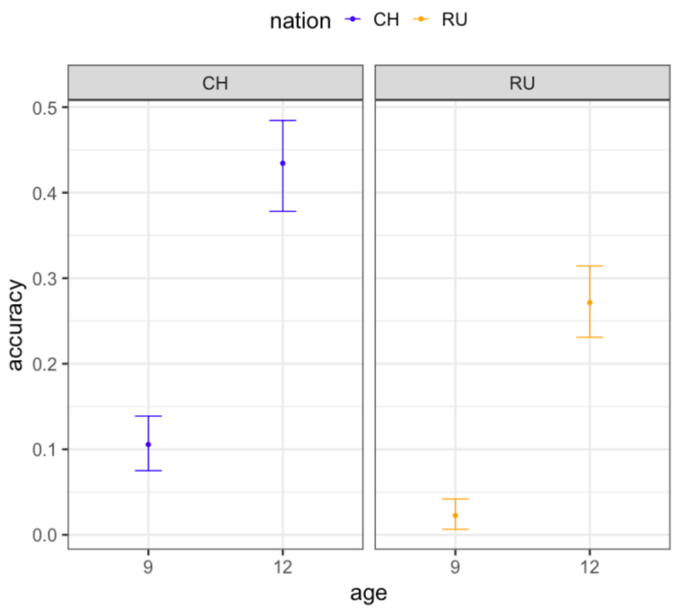


Figure 9.1: Mean accuracy on 3SG-agreement per L1 and age group

Results of the final model showed three significant main effects, and a significant three-way interaction (c.f. Table 9.2). The first significant main effect concerned L1, $\chi^2(1)=17.48$, $p<.001$, with significantly lower accuracy in 3SG marker for the Russian group ($M=0.017$, *SE*=0.59) compared to the Chinese learners (*M*=0.072, *SE*=0.52). The second significant main effect was age, $\chi^2(1)=16.75$, $p<.001$, with 9-year-olds (*M*=0.004, *SE*=0.61) performing significantly worse than 12-year-olds (*M*=0.19, *SE*=0.41). There was also a significant main effect of vocabulary $\chi^2(1)=25.41$, $p<.001$, with higher scores on it entailing higher accuracy. Finally, there was a significant three-way interaction between L1, age, and vocabulary, $\chi^2(1)=15.54$, $p<.05$. No other effects reached significance.

To further understand the three-way interaction, I divided the data by L1. The interaction between age and vocabulary was significant for the Chinese 9-year-olds, $\chi^2(1)=16.80$, $p<.001$, but was not found to be significant for the Chinese 12-year-olds $\chi^2(1)=1.36$, $p=.48$. In contrast, the interaction between vocabulary and age was significant for the Russian 12-year-olds, $\chi^2(1)=23.39$, $p<.001$, but not for the Russian 9-year-olds, $\chi^2(1)=0.13$, $p=1$.

Table 9.2: Optimal model with accuracy on 3SG-agreement as the dependent variable; significant main effects and interactions

	Chisq	**Df**	**Pr(>Chisq)**
nation	17.4815	1	2.901e-05 ***
age	16.7566	1	4.249e-05 ***
zvoc	25.4111	1	4.632e-07 ***
nation:age	1.7322	1	0.1881
nation:zvoc	2.0544	1	0.1518
age:zvoc	0.0967	1	0.7558
nation:age:zvoc	15.5453	1	8.055e-05 ***

*p < .05. **p < .01. ***p <. 001.

These findings suggest that accuracy in 3SG-agreement is significantly higher for older learners than younger learners, that Chinese learners do significantly better than Russian learners in 3SG, and that vocabulary interacts with age and L1 and is predictive of performance on 3SG; for Chinese 9-year-olds higher vocabulary score means higher accuracy in 3SG but vocabulary is not predictive of the Chinese 12-year-olds' performance; for Russians, the reverse is true, 12-year-olds' higher vocabulary also means higher accuracy in 3SG but the vocabulary score is not predictive of Russian 9-year-olds' performance. In other words, vocabulary explains accuracy for Chinese 9-year-olds and Russian 12-year-olds but not for the other two groups.

9.3.2 Binomial logistic regression for TEGI Past regular

Figure 9.2 offers a schematic view of the mean accuracy results on regular past for each age and L1 group.

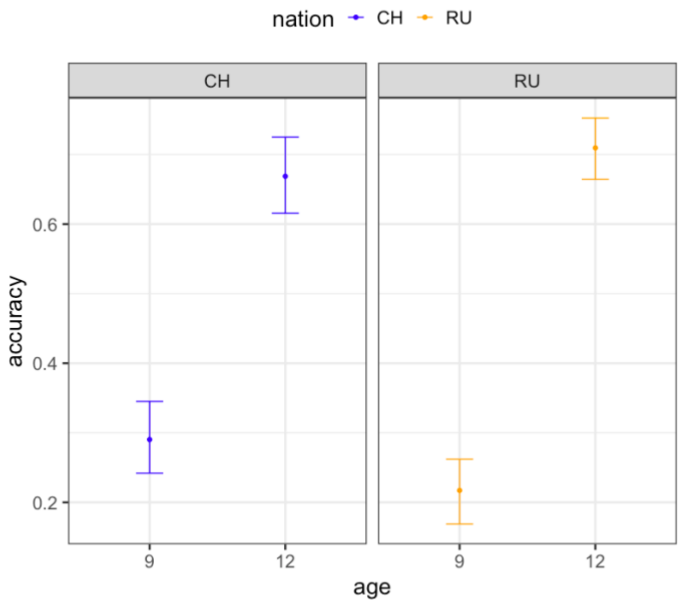


Figure 9.2: Mean accuracy on regular past per L1 and age group

Results of the optimal model -as Table 9.3 shows- revealed two significant main effects and a significant three-way interaction. Specifically, there was a significant main effect of age $\chi^2(1)=23.89$, $p<.001$, with 12-year-olds ($M=0.84$, $SE=0.47$) outperforming 9-year-olds ($M=0.04$, $SE=0.60$). Vocabulary also demonstrated a significant effect, $\chi^2(1)=19.93$, $p<.001$ with higher vocabulary scores entailing higher accuracy. Finally, there was a significant three-way interaction between L1, age, and vocabulary score, $\chi^2(1)=10.37$, $p<.01$.

Table 9.3: Optimal model with accuracy on regular past as the dependent variable; significant main effects and interactions

	Chisq	**Df**	**Pr(>Chisq)**
nation	1.8363	1	0.175391
age	23.8950	1	1.017e-06 ***
zvoc	19.9340	1	8.016e-06 ***
nation:age	0.1345	1	0.713852
nation:zvoc	3.7652	1	0.052331 .
age:zvoc	0.0171	1	0.896061
nation:age:zvoc	10.3725	1	0.001279 **

*p < .05. **p < .01. ***p <. 001.

To further explore the three-way interaction, I again divided the data by L1. The interaction between age and vocabulary was significant for both Chinese groups; thus, vocabulary was predictive of accuracy for 9-year-olds, $\chi^2(1)=18.07$, $p<.001$, and for 12-year-olds, $\chi^2(1)=7.14$, $p<.05$. This interaction between vocabulary and age was also significant for the Russian cohort but only for the 12-year-olds, $\chi^2(1)=9.27$, $p<.01$ while it was not significant for the 9-year-olds. Figure 9.3 illustrates the three-way interaction visually.

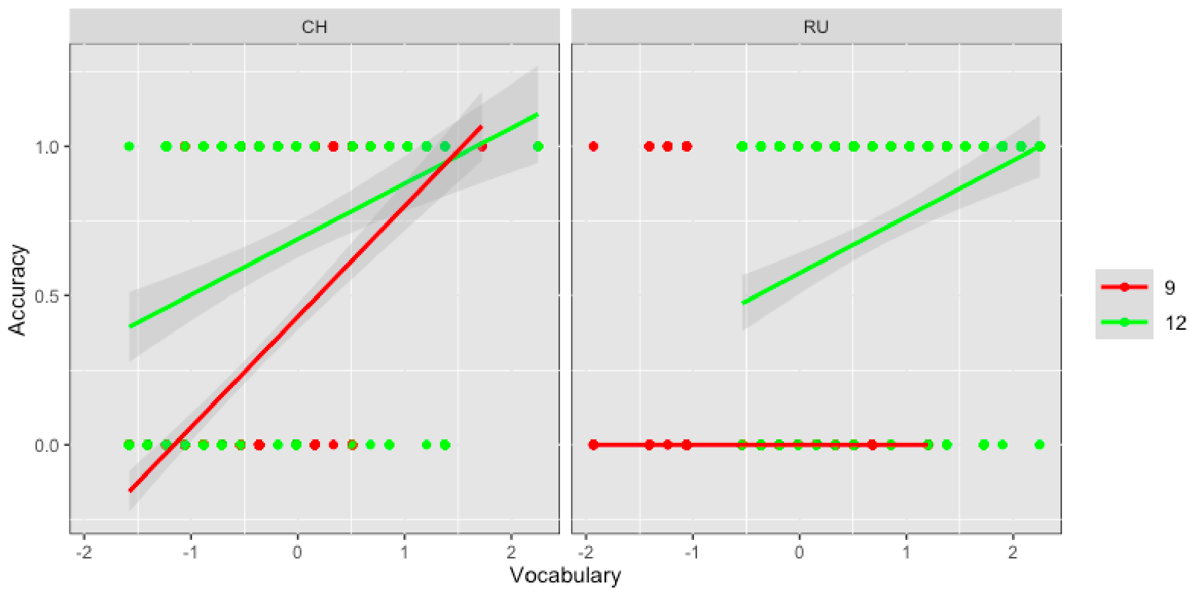


Figure 9.3: The three-way interaction between L1, age, and vocabulary score

These findings show that the two L1 cohorts did not differ in past regular as there was no significant main effect of L1. Accuracy in regular past tense is significantly higher for older learners than younger learners, that is older children outperformed the younger ones. Finally, vocabulary interacts with age and is predictive of accuracy on regular past verbs for all groups except the Russian 9-year-olds.

9.3.3 Binomial logistic regression for TEGI (irregular) Past

Figure 9.4 displays mean accuracy on irregular past for each L1 and age group.

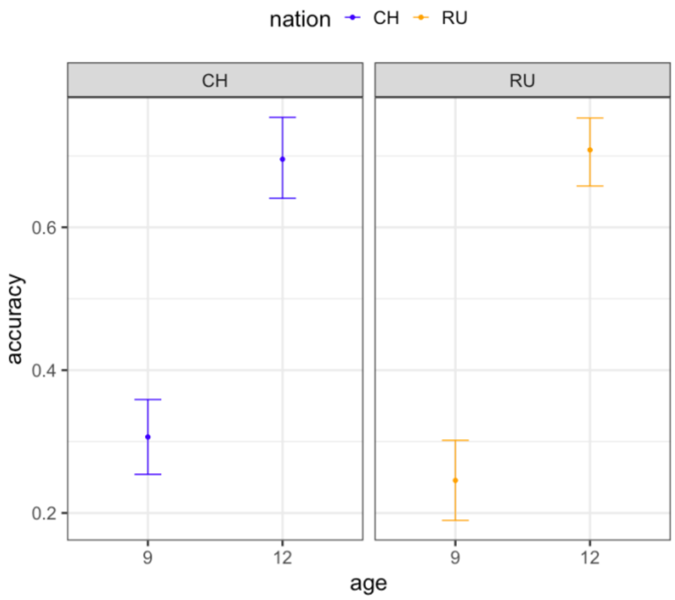


Figure 9.4: Mean accuracy on irregular past per L1 and age group

The optimal model (c.f. Table 9.4) demonstrated three significant main effects; a significant main effect of L1 $\chi^2(1)=4.43$, $p<.05$ with Russians (M=0.52, SE=0.50) performing better than Chinese (M=0.47, SE=0.53) on irregular verbs. Age was also a significant predictor of accuracy in past irregular, $\chi^2(1)=20.00$, $p<.001$, with older children (M=0.88, SE=0.50) doing significantly better than younger ones (M=0.07, SE=0.57). Finally, there was a significant effect of vocabulary, $\chi^2(1)=27.09$, $p<.001$.

Consistent with previous results, 9-year-olds do significantly worse than 12-year-olds. In contrast, Russian children do significantly better than Chinese learners. Finally, vocabulary is predictive of children's performance as the higher their vocabulary score, the better their performance on past irregular.

Table 9.4: Optimal model with accuracy on irregular past as the dependent variable; significant main effects and interactions

	Chisq	Df	Pr(>Chisq)
nation	4.4383	1	0.03514 *
age	20.0045	1	7.726e-06 ***
zvoc	27.0955	1	1.936e-07 ***

*p < .05. **p < .01. ***p <. 001.

9.3.4 Binomial logistic regression for accuracy on past tense as an effect of verb regularity

Figure 9.5 presents the accuracy results on past tense for each age and L1 group.

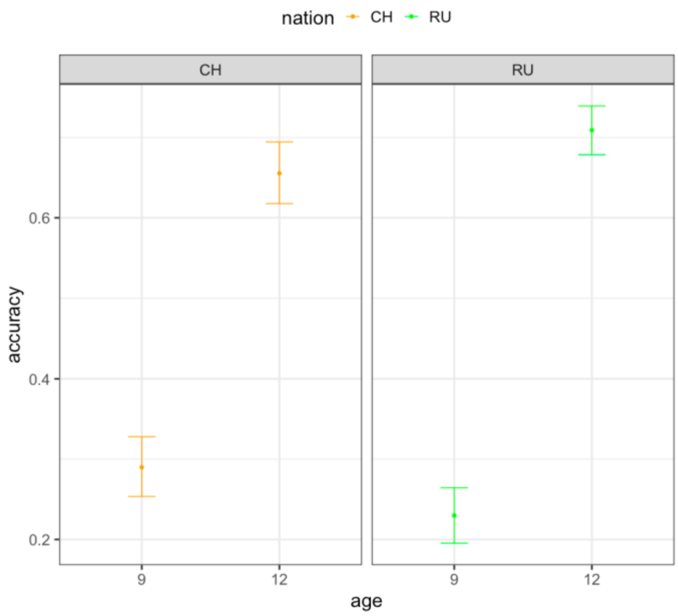


Figure 9.5: Mean accuracy on past tense (both regular and irregular verbs) per L1 and age group

Recall that the aim of this analysis was to find out whether the regularity status of past tense verbs would affect accuracy. Results of the final model -as Table 9.5 displays- showed only one significant main effect; age, $\chi^2(1)= 42.55$, $p<.001$, with significantly lower accuracy in past tense for the younger children (*M*=0.07, *SE*=0.48) compared to the older learners (*M*=0.80, *SE*=0.40). No other effects reached significance.

Table 9.5: Optimal model with accuracy on past tense as the dependent variable; significant main effects and interactions

	Chisq	Df	Pr(>Chisq)
age	42.557	1	6.865e-11 ***

This entails that accuracy could not be predicted by whether a verb was regular or irregular and this holds across ages and L1s.

9.3.5 Binomial logistic regression for the asymmetry in the acquisition of features

Figures 9.6-9.7 offer the mean accuracy on the two morphemes for each L1 and age group for inspection.

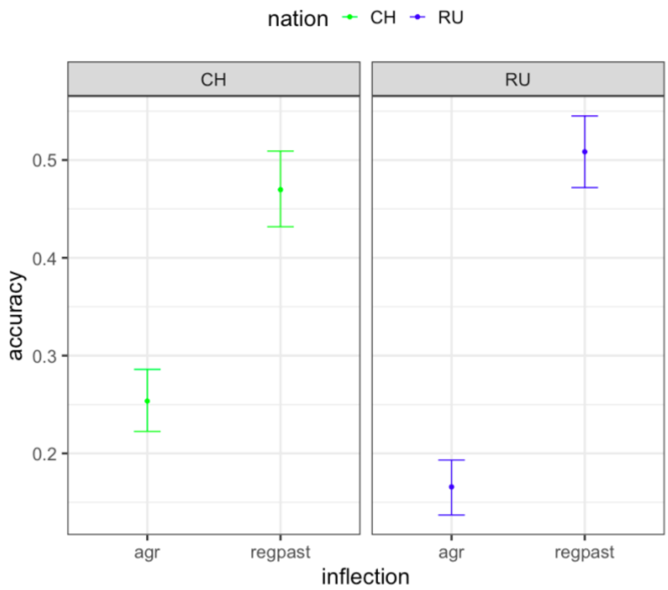


Figure 9.6: Mean accuracy on 3SG-agreement vs regular past per L1

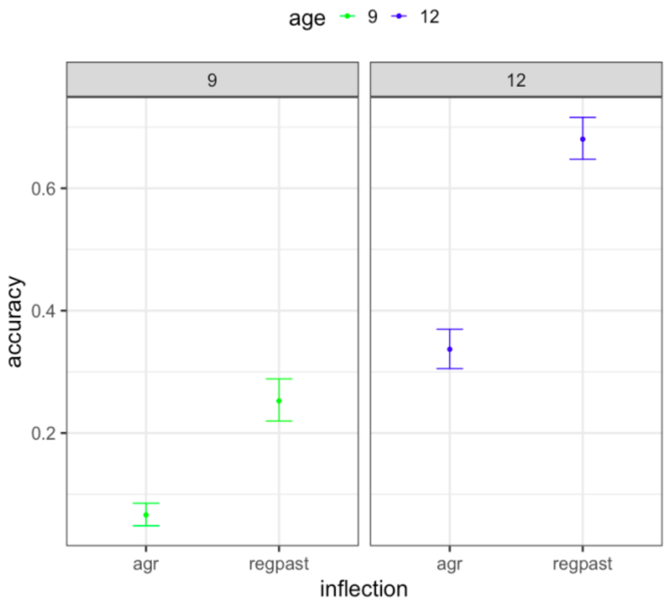


Figure 9.7: Mean accuracy on 3SG-agreement vs regular past per age groups

The final model showed two significant main effects and two significant two-way interactions as shown in Table 9.6. The first significant main effect was age $\chi^2(1)=48.6$, $p<.001$, with older learners ($M=0.48$, $SE=0.22$) outperforming the younger ones ($M=0.03$, $SE=0.44$). The second significant main effect was tense, $\chi^2(1)=239.7$, $p<.001$ with accuracy on regular past ($M=0.49$, $SE=0.37$) being significantly higher than accuracy on 3SG-agreement ($M=0.10$, $SE=0.38$). There was also a significant two-way interaction between L1 and tense, $\chi^2(1)=14.51$, $p<.001$, showing that both L1 groups do better in past tense marking than in agreement marking (Chinese children: past ($M=0.44$, $SE=0.53$), agreement ($M=0.12$, $SE=0.53$); Russian children: past ($M=0.55$, $SE=0.45$), agreement ($M=0.08$, $SE=0.46$)), but the effect was significantly stronger in Russian than in Chinese. Finally, a significant two-way interaction between tense and age was found, $\chi^2(1)=7.9$, $p<.01$; both 12-year-olds and 9-year-olds performed significantly higher on past tense (12y.o.: $M=0.71$, $SE=0.25$, 9y.o.: $M=0.06$, $SE=0.52$) than 3SG-agreement (12y.o.: $M=0.27$, $SE=0.25$, 9y.o.: $M=0.004$, $SE=0.57$) but 12-year-olds outperformed 9-year-olds in this respect.

These results suggest that in all cases, that is across age groups and L1s, children did significantly better in past (regular) than agreement.

Table 9.6: Optimal model with accuracy in both 3SG-agreement and regular past verbs as the dependent variable; significant main effects and interactions

	Chisq	**Df**	**Pr(>Chisq)**
nation	0.7766	1	0.3781802
age	48.5957	1	3.146e-12 ***
tense	239.7162	1	< 2.2e-16 ***
nation:tense	14.5115	1	0.0001393 ***
tense:age	7.9051	1	0.0049296 **

*p < .05. **p < .01. ***p <. 001.

9.3.6 Summary of statistical results for TEGI tasks

- Accuracy on 3SG-agreement is predicted by:
    - age; 12-year-olds did significantly better than the 9-year-olds
    - L1; Chinese did significantly better than Russians
    - vocabulary; Renfrew vocabulary score explained accuracy, but only for Chinese 9-year-olds and Russian 12-year-olds.

- Accuracy on past regular is predicted by:
    - age; 12-year-olds performed significantly better than 9-year-olds
    - vocabulary; the higher the Renfrew vocabulary score, the better the performance on regular past
    - vocabulary; predictive of performance for all groups except the Russian 9-year-olds

- Accuracy on past irregular is predicted by:
    - age; 12-year-olds did significantly better than 9-year-olds
    - L1; Russians performed significantly better than Chinese
    - vocabulary; higher scores in the vocabulary task meant higher accuracy in irregular past.

- Accuracy on agreement vs past (regular):
    - accuracy is significantly higher in past (regular) than agreement across ages and L1s.

9.4 Statistical analyses in narrative production

To address the question of which factors predict accuracy/overt finiteness marking in narrative production, I ran a similar analysis as in the case of TEGI with accuracy as the dependent variable and the same independent variables but also including the narrative length and the verb type.

Apart from 3 Chinese children (CH_GA_35, CH_GB_23, CH_GB_25) who were excluded from the analysis due to missing academic hours, one more child from the Russian cohort (RU_GB_37) was excluded as it was an outlier in the narrative length measure[31].

In analyses on TEGI, the only measure of proficiency I had was the Renfrew Vocabulary score. However, after analyzing narrative productions, I calculated two additional measures, that is, the narrative length (i.e. the number of clauses) and the lexical diversity (i.e. type-token ratio). I take these two measures as a proxy for proficiency since research on linguistic complexity in SLA shows a systematic correlation between linguistic complexity and proficiency (Norris & Ortega, 2009). In order to decide on which of these variables to include in the model and avoid collinearity, I ran correlational analyses between them.

Table 9.7 shows the Pearson's coefficients of the correlational analyses carried out between pairs of continuous predictor variables as well as the p-values. All are significantly correlated showing a positive correlation. The Renfrew vocabulary score is strongly correlated with lexical diversity (type-token ratio). In addition, both measures showed a discrepancy between older groups but no difference between younger groups. Due to their correlation, only one can be included in the analysis; I included the Renfrew score for consistency with the previous analyses for TEGI tests. The Renfrew score is also correlated with the narrative length measure, but the size effect of this correlation is not as strong. Moreover, the narrative length measure discriminates between the younger children showing that Russians produced a significant lower number of clauses than Chinese and can thus capture a complementary aspect of proficiency. Thus, both of these measures (i.e. the Renfrew vocabulary score and the narrative length score) will be used as a proxy for proficiency.

[31] The number of clauses of this child was extremely small and plotting of the data showed it as an outlier.

Table 9.7: Correlational analyses between continuous predictor variables: Pearson's correlation coefficients and p-values

Variables	Correlation coefficient	p-value
Renfrew Voc score - Narrative length	0.403793 (medium effect size)	6.258e-07
Lexical diversity - Narrative length	0.2875154 (medium effect size)	0.0005216
Renfrew Voc score – Lexical diversity	0.5929414 (large effect size)	7.579e-15

Figure 9.8 displays the overall (i.e. includes all verb types: copula *be*, auxiliary *be*, main thematic verb) accuracy results for each age and L1 group.

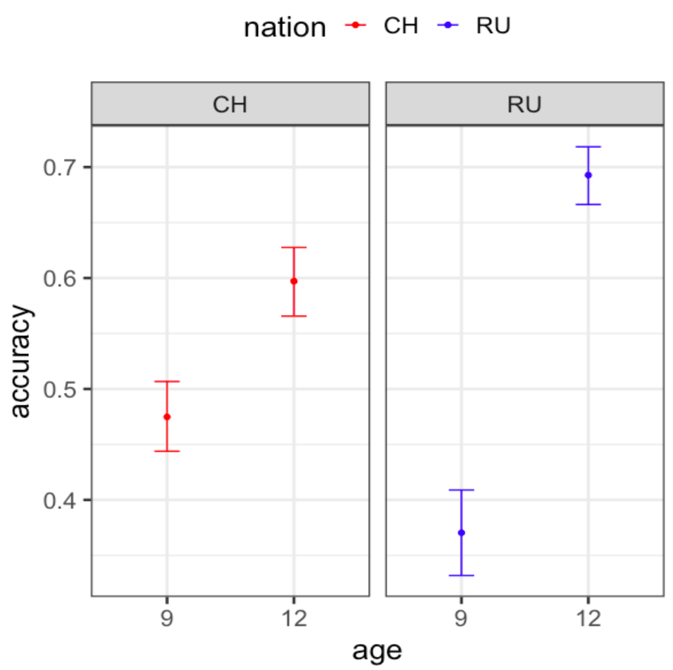


Figure 9.8: Mean overall accuracy on overt finiteness marking in narrative production per age and L1 group

The initial full fixed-effects structure thus included two three-way interactions: L1, age, and length of narratives (i.e. nation x age x znar_length[32]) and L1, age, and academic hours of attendance at EF (i.e. nation x age x zhours) and two main effects: the Renfrew vocabulary

[32] Recall that continuous variables such as the length of the narratives and the hours were z-score transformed (standardised), hence znar_length or zhours.

score, and verb type. Finally, a random intercept for participant was applied. More complex random structures caused convergence issues.

The optimal model – as Table 9.8 shows- demonstrated four significant main effects and two significant two-way interactions. The first main effect that was significant was age, $\chi^2(1)=13.35$, $p<.001$, with significantly higher accuracy for the older learners ($M=0.67$, $SE=0.12$) compared to the younger ones ($M=0.39$, $SE=0.12$). The second significant main effect was the Renfrew vocabulary score $\chi^2(1)=34.31$, $p<.001$, that is children with higher scores on the Renfrew Vocabulary task were more accurate than those having lower scores. The third significant main effect was children's narrative length, $\chi^2(1)=32.96$, $p<.001$; the longer the stories the children produced, the more accurate they were in terms of verb morphology. The fourth significant main effect was the verb type, $\chi^2(2)=505.94$, $p<.001$, with significantly higher accuracy on copula *be* ($M=0.93$, $SE=0.19$) than auxiliary *be* ($M=0.78$, $SE=0.17$) and main thematic verb ($M=0.37$, $SE= 0.12$). Auxiliary *be* in turn showed higher accuracy than main thematic verbs. Finally, there were two significant two-way interactions; the first one between L1 and narrative length, $\chi^2(1)=4.65$, $p<.05$. Narrative length was a significant predictor of accuracy for both L1 cohorts but more so for Russians as Figure 9.9 displays. The second significant two-way interaction was between age and academic hours of attendance at EF, $\chi^2(1)=7.37$, $p<.01$; specifically, the significant interaction concerns the 12-year-olds showing a negative association with academic hours as predictive of verb accuracy as Figure 9.10 illustrates – for 9-year-olds although a positive association is observed this is not significant.

Table 9.8: Optimal model with overall verb accuracy as the dependent variable; significant main effects and interactions

	Chisq	Df	p-value
nation	0.6311	1	0.4269361
age	13.3579	1	0.0002573 ***
zvoc	34.3143	1	4.689e-09 ***
znar_length	32.9632	1	9.392e-09 ***
verbtype	505.9473	2	< 2.2e-16 ***
zhours	0.9686	1	0.3250322
nation:znar_length	4.6576	1	0.0309163 *
age:zhours	7.3719	1	0.0066250 **

*p < .05. **p < .01. ***p <. 001.

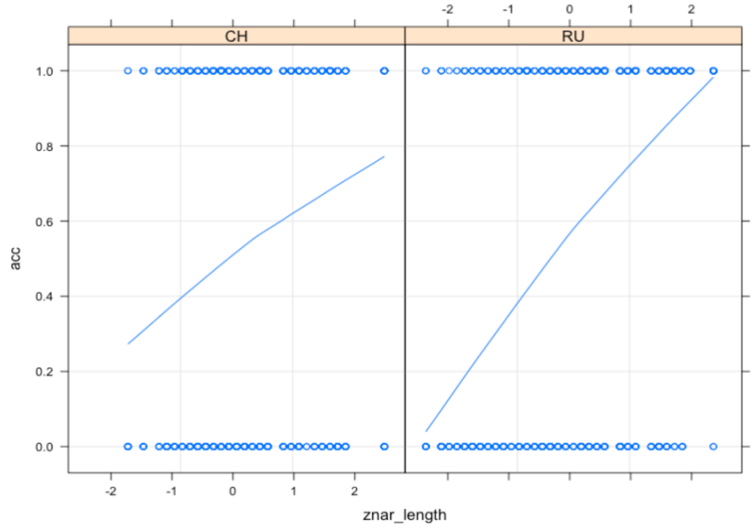


Figure 9.9: Two-way interaction of L1 and narrative length as predictors of verb form accuracy

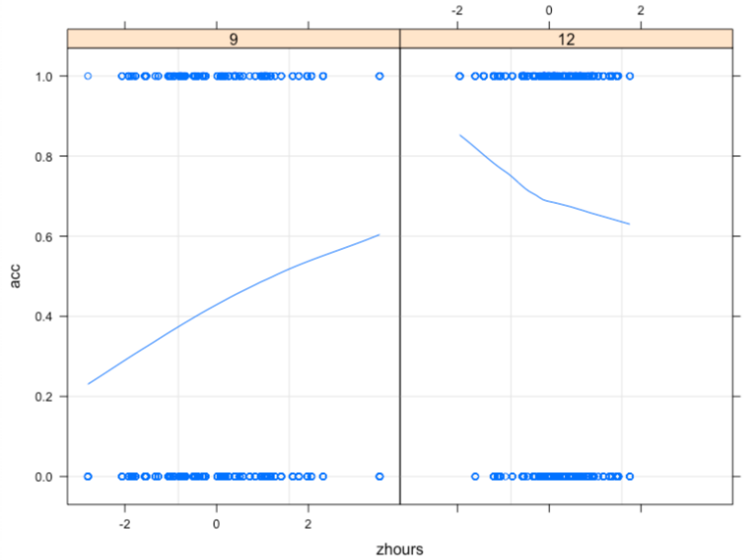


Figure 9.10: Two-way interaction of age and academic hours as predictors of verb form accuracy

These results suggest that 1) verb form accuracy increases with age, that is, older children outperform younger ones, 2) accuracy differs depending on the verb type, that is, children are significantly more accurate on copula *be* than auxiliary *be*, and more accurate on auxiliary *be* than the main thematic verb inflection, 3) accuracy increases with proficiency; the length of the narratives as well as the vocabulary size play a significant role in predicting accuracy, and 4) verb form accuracy is not systematically dependent on academic hours of attendance as for older children there is a significant negative association while for younger ones the association is positive although non-significant.

Following this analysis, I decided to further look at what predicts accuracy for each verb type separately; copula *be*, auxiliary *be*, and main thematic verbs and whether the same predictors were significant in each case.

9.4.1 Mixed-effects binomial logistic regression for accuracy on copula *be*

Figure 9.11 illustrates the mean accuracy results on copula *be* for each age and L1 group.

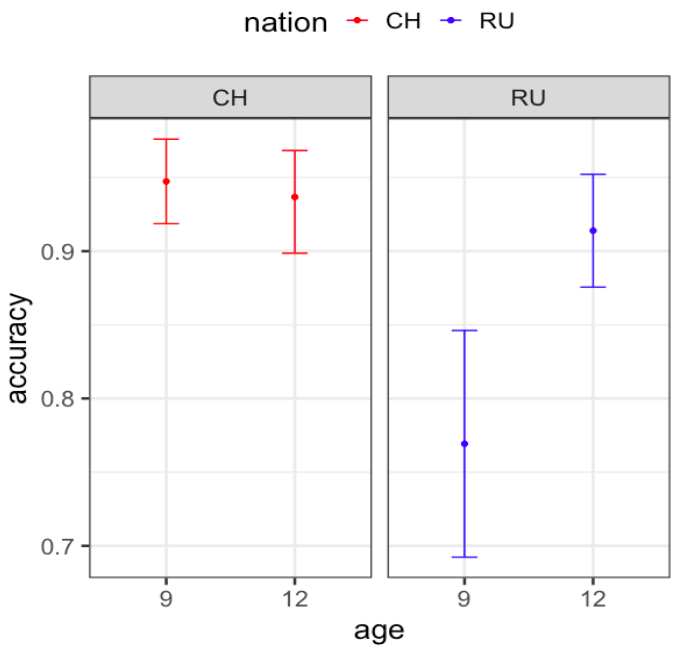


Figure 9.11: Mean accuracy on copula be per age and L1 groups

Accuracy on copula *be* was the dependent variable in this analysis and was modelled as a function of the fixed structure involving two three-way interactions and two main effects. Specifically, the first three-way interaction involved the L1, the age, and the narrative length measure (nation x age x znar_length) and the second one included the L1, the age and the academic hours of attendance in EF (nation x age x zhours). I also added two main effects; the Renfrew vocabulary score and the number (3SG vs 3PL). I finally applied a random intercept for participant. More complex structures raised convergence issues.

The final model revealed three significant main effects, one significant two-way interaction and a significant three-way interaction (c.f. Table 9.9). The first significant main effect was L1, $\chi^2(1)=4.39$, $p<.05$, with Chinese learners ($M=0.96$, $SE= 0.37$) outperforming the Russian learners ($M=0.91$, $SE= 0.30$). The second significant main effect was number,

$\chi^2(1)=26.99$, $p<.001$, with accuracy on the 3SG condition ($M=0.97$, $SE=0.38$) being significantly higher than on the 3PL condition ($M=0.84$, $SE=0.37$). The third significant main effect was the narrative length $\chi^2(1)=4.46$, $p<.05$. Finally, there was a significant three-way interaction between L1, age, and narrative length. To further explore this interaction, I divided the data by L1. There was no significant main effect or interaction for the Chinese group while there was a significant main effect of narrative length, $\chi^2(1) = 12.15$, $p < .001$, and a significant two-way interaction between age and narrative length, $\chi^2(1)=5.70$, $p<.05$) for the Russian cohort. Specifically, narrative length was a highly significant predictor of accuracy on copula *be* for the Russian group while the significant two-way interaction concerned only the 9-year-olds; Russian 9-year-olds who had higher narrative length did better on copula *be* but it was not significant for the 12-year-olds (perhaps because they already had high accuracy).

Table 9.9: Optimal model with accuracy on copula be as the dependent variable; significant main effects and interactions

	Chisq	Df	p-value
nation	4.3933	1	0.036080 *
age	0.0370	1	0.847400
znar_length	4.4615	1	0.034668 *
number	26.9983	1	2.036e-07 ***
nation:age	0.0050	1	0.943518
nation:znar_length	8.6826	1	0.003213 **
age:znar_length	0.7861	1	0.375276
nation:age:znar_length	8.0711	1	0.004498 **

*p < .05. **p < .01. ***p <. 001.

In a nutshell, narrative length predicts accuracy on copula *be* only for the Russian 9-year-olds as Figure 9.12 illustrates. For older Russians as well as for the Chinese groups, performance was much higher, perhaps because they were all above a threshold of exposure and proficiency in English. The fact that Russians appear to do worse is in fact only because younger Russian children have very low proficiency and their performance is considerably lower than all the other groups.

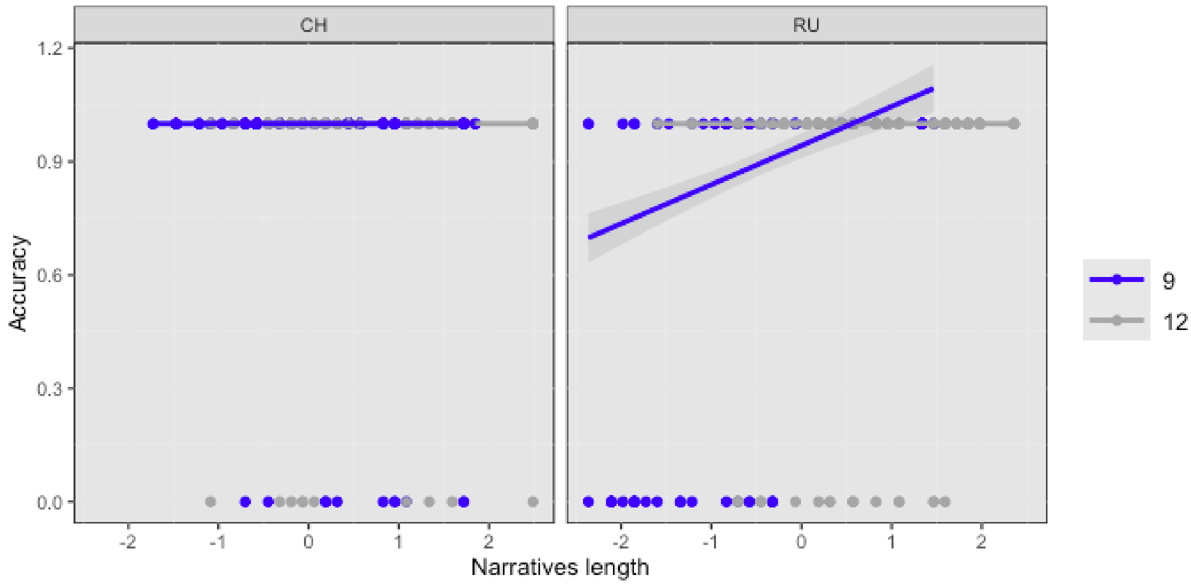


Figure 9.12: The three-way interaction between L1, age and narrative length; narrative length was only predictive for the Russian 9-year-olds

9.4.2 Mixed-effects binomial logistic regression for accuracy on auxiliary *be*

Figure 9.13 displays the accuracy results on auxiliary *be* for each L1 and age group.

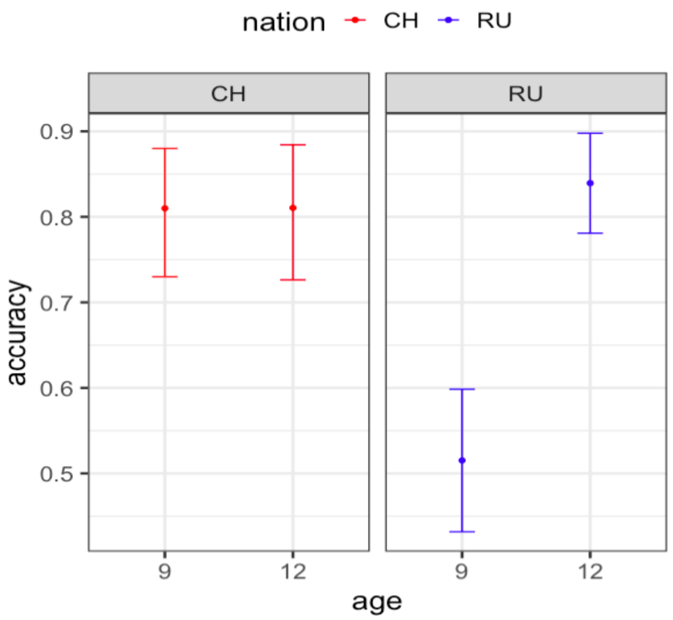


Figure 9.13: Mean accuracy on auxiliary be per L1 and age group

To find out what predicts accuracy on auxiliary *be*, accuracy was modelled as a function of the same structure as described in the previous model with the same random structure applied.

The optimal model demonstrated two significant main effects as shown in Table 9.10. The first significant main effect was narrative length, $\chi^2(1)=50.08$, $p<.001$, showing that the longer the stories the children produced, the higher the accuracy on auxiliary *be*. The second significant main effect was number, $\chi^2(1)=4.38$, $p<.05$, with accuracy on the 3SG condition ($M=0.80$, $SE=0.17$). being significantly higher than accuracy on the 3PL condition ($M=0.61$, $SE= 0.44$). No other effect or interaction reached significance.

Table 9.10: Optimal model with accuracy on auxiliary be as the dependent variable; significant main effects and interactions

	Chisq	Df	p-value
number	4.3837	1	0.03628 *
znar_length	50.0843	1	1.473e-12 ***

*p < .05. **p < .01. ***p <. 001.

9.4.3 Mixed-effects binomial logistic regression for accuracy on main thematic verbs

Let us finally turn to main thematic verbs. Figure 9.14 displays the accuracy results by L1 and age groups.

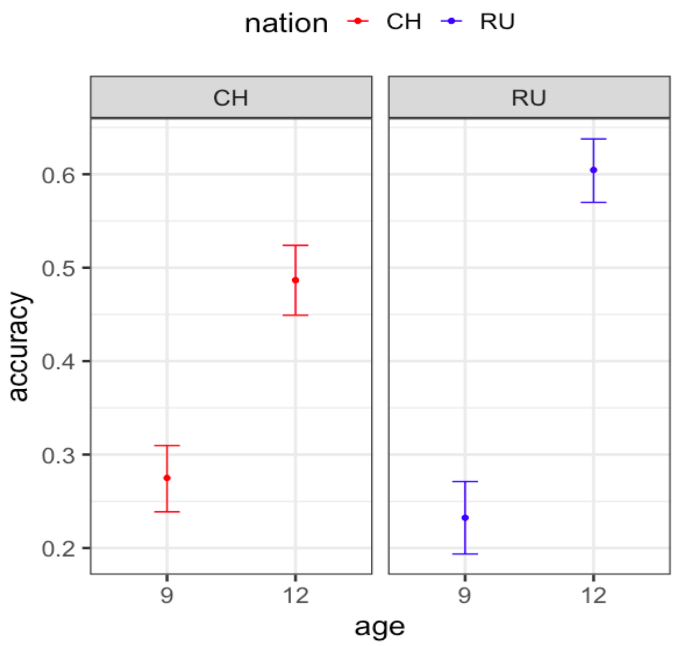


Figure 9.14: Mean accuracy on main thematic verbs per L1 and age group

Accuracy on main verbs was the dependent variable in the last analysis. The argument structure of the fixed effects was the same as described above but number was not included in this analysis -only 3SG verbs were considered. There was no point in considering plural contexts where bare forms are legitimate when we are looking into inflection. The structure I applied involved the two three-way interactions and a main effect (i.e. L1 x age x narrative length; L1 x age x hours and vocabulary) while I also applied a random intercept for participant.

The optimal model showed three significant main effects and a significant two-way interaction – illustrated in Table 9.11. The first significant main effect was age, $\chi^2(1)=17.58$, $p<.001$, with older children ($M=0.56$, $SE=0.15$) outperforming the younger ones ($M=0.19$, $SE=0.17$). The second significant main effect was narrative length, $\chi^2(1)=16.82$, $p<.001$, and the third one the Renfrew vocabulary score, $\chi^2(1)=37.77$, $p<.001$. Both showed that the longer the stories and the higher the Renfrew score, the higher the accuracy on inflection on main thematic verbs. Finally, there was a significant two-way interaction between age and academic hours of attendance in EF, $\chi^2(1)=7.99$, $p<.01$. Post-hoc analyses using the Bonferroni method showed that this interaction was significant only for the 12-year-old children showing a negative effect; the more the hours they had attended, the lesser their accuracy. This interaction was not significant for the 9-year-olds meaning that hours did not explain accuracy. Figure 9.15 illustrates this effect of hours on accuracy concerning only the 12-year-olds.

Table 9.11: Optimal model with accuracy on main thematic verbs as the dependent variable; significant main effects and interactions

	Chisq	Df	p-value
age	17.5896	1	2.741e-05 ***
znar_length	16.8256	1	4.098e-05 ***
zhours	0.8806	1	0.348032
zvoc	37.7787	1	7.924e-10 ***
age:zhours	7.9906	1	0.004702 **

*p < .05. **p < .01. ***p <. 001.

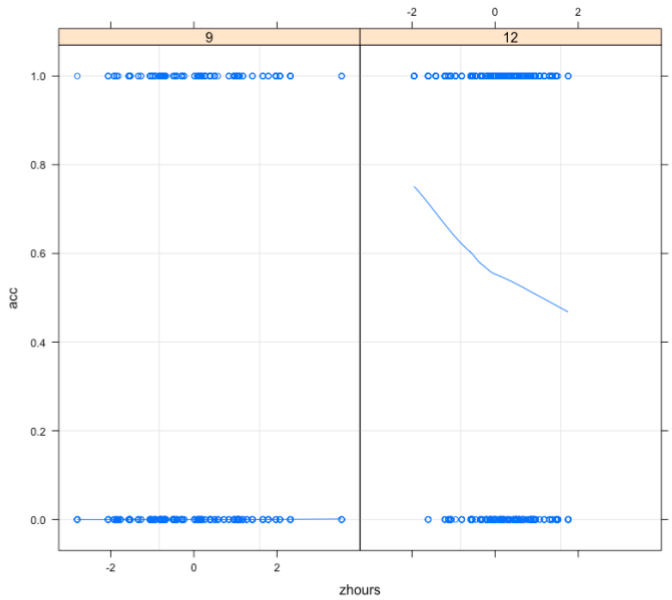


Figure 9.15: Two-way interaction of age and academic hours as predictors of verb form accuracy

9.4.4 Summary of statistical results for the narratives

- Overall accuracy/overt finiteness marking is predicted by:
    - age; 12-year-olds outperformed the 9-year-olds
    - verb type; accuracy on cop be > aux be > main thematic verb
    - narrative length; the higher the number of clauses or the longer the stories, the higher the verb form accuracy
    - academic hours; for older children only accuracy showed a negative association with increased academic hours

- Accuracy on copula *be* is predicted by:
    - L1; Chinese did better than Russians
    - number; 3SG showed higher accuracy than 3PL
    - narrative length; significant for the Russian 9-year-olds only

- Accuracy on auxiliary *be* is predicted by:
    - number; 3SG showed higher accuracy than 3PL
    - narrative length; the longer the stories, the higher the verb form accuracy

- Accuracy on main thematic verbs is predicted by:
    - age; older children outperformed younger ones
    - narrative length; the longer the stories, the higher the verb form accuracy
    - Renfrew vocabulary score; the higher the vocabulary score, the higher the verb form accuracy of main thematic verbs
    - academic hours; negative effect of increased number of hours on accuracy on main thematic verbs for 12-year-olds.

9.5 Summary of results

In this chapter, I presented the statistical analyses concerning the predictors of verb form accuracy in two types of tasks; the TEGI tasks and a narrative production task. Let us now see the picture emerging from the findings in both tasks that describe the important factors on the acquisition of finiteness by children. These are:

- *Age*

Age proved to be a significant predictor of verb form accuracy in most of the analyses in both tasks. In all analyses of main thematic verb accuracy relevant to the TEGI tasks as well as to the narrative task it was found that 12-year-olds did consistently and significantly better than 9-year-olds. In contrast, the accuracy on copula and auxiliary *be* (narrative task) showed no effect of age.

- *L1*

L1 was found to affect accuracy but not in the same direction in all morphemes. Specifically, while Chinese did better on 3SG-agreement of the TEGI task, there was no significant difference between the two L1 cohorts for past regular, while in past irregular, Russians outperformed Chinese learners. With respect to accuracy on copula *be* in narrative production, Chinese again surpassed Russian learners while no difference between the two L1 cohorts was revealed when analyzing accuracy on auxiliary *be* and main thematic verbs.

A short note here is that L1 effects do seem in some cases to be ‘masked’. There is a proficiency issue arising when considering Russian 9-year-old children. Their proficiency as assessed through their narrative length as well as considering their average CEFR class level is lower than the proficiency of their Chinese counterparts, even if they have followed 5 years of instruction and the same curriculum with the Chinese children. Inspecting all graphs, younger Chinese group is at a post-beginner level while the Russian is still at beginner level; this discrepancy in proficiency disappears in older children, but for younger groups it means we cannot properly assess L1 effects because the two groups are in effect of different proficiency. This difference between the proficiency levels of the two younger groups seem to make a huge difference when considering overall L1 effects.

- *Renfrew Vocabulary score*

The Renfrew vocabulary score was found to be a significant predictor of accuracy on 3SG-agreement, past regular and irregular analyses of the TEGI task as well as in analysis examining predictors of accuracy of main thematic verbs in narratives.

There were also some interesting findings about its specific role. First, vocabulary size could explain accuracy on 3SG agreement in the TEGI task only for Chinese 9-year-olds and Russian 12-year-olds. Younger Chinese with higher proficiency show higher accuracy, but for older Chinese children, higher proficiency does not translate into higher accuracy. In other words, higher proficiency as captured by the vocabulary score does not mean higher accuracy on 3SG-agreement for Chinese 12-year-olds. Considering the Russian cohort, we see the reverse; the Renfrew vocabulary score does not predict accuracy on 3SG-agreement for younger children while older Russians do improve as proficiency increases. We come back to this point in the discussion chapter.

Considering results of the statistical analysis on regular past accuracy of TEGI task, the Renfrew Vocabulary score could explain accuracy for all groups except for Russian 9-year-olds. This shows that we need to reconsider children’s proficiency in the discussion chapter before proceeding to group comparisons.

- *Narrative length*

This measure proved to be a significant predictor of children's verb form accuracy. It accounted for variance in accuracy on auxiliary *be* and main thematic verbs; the longer the children's narratives would be, the higher the verb form accuracy concerning these morphemes.

Interestingly, narrative length was also a significant predictor of accuracy on copula *be* but this concerned especially the Russian 9-year-olds and none of the other groups. It is clear that with higher proficiency, Russian 9-year-olds do better on copula *be* but they still have not reached at least a threshold of proficiency as all the other groups to have learnt the copula *be*.

- *Academic hours*

Academic hours were used as a measure of exposure/input children had received and was included to control for such effects. Academic hours did not prove significant in any of the TEGI analyses. However, in the narratives we found that increasing academic hours predicted lower main thematic verb form accuracy. This may be surprising at first sight. I speculate that this could be the case if the children have not acquired the relevant morphemes and under exposure to new grammatical phenomena and structures, they go through a U-shape in their development. For younger children academic hours had no negative or positive impact in any of the tasks.

- *Verb type*

In analyses of the narrative data, I also included the verb type as a predictor of accuracy. It was found to be a significant predictor of it revealing that accuracy was significantly higher for copula *be* than auxiliary *be*, and on auxiliary *be* higher than main thematic verbs.

- *Number*

In analysis of accuracy on copula and auxiliary *be*, another predictor variable was number as a difference was attested between 3SG and 3PL contexts when presenting the data in the previous chapter. In both cases, 3SG was found to be a significant higher predictor of accuracy compared to the 3PL condition.

- *Inflection (3SG-agreement vs regular past)*

Another analysis on TEGI looked at whether accuracy differed significantly between 3SG-agreement and regular past tense. The inflectional morpheme past was found to be easier than 3SG-agreement, meaning that it showed higher accuracy, across ages and L1s.

I did not consider accuracy as a function of inflection (3SG vs past) in narratives because there were no obligatory contexts and such a predictor would not really tell us anything. However, considering raw results and percentages it is clear that children, with the exception of Chinese 9-year-olds, used much more past tense than 3SG-agreement markers (see Section 7.5.4).

CHAPTER 10

DISCUSSION

10.0 Introduction

In this chapter, I consider the results presented in the last chapters in connection to the questions and hypotheses motivating the studies as presented in Chapter 5.

I begin in Section 10.1 with the question of proficiency of the two L1 cohorts which is crucial for comparing and interpreting children's performance. In Section 10.2, I briefly present the L2 development of children with different ages of onset in EF. I then discuss the results in the light of the relevant hypotheses and theories. I start in Section 10.3 with discussing whether younger children and older children show the same performance or not as a result of their different ages of onset. I then turn in Section 10.4 to the issue of L2 children's performance as an effect of limited input. After I show that input is not the determinant factor for children's performance and error patterns, I return -in section 10.5- to Meisel's hypothesis (2009) to consider if children older than 4 pattern with L2 adults in the domain of inflectional morphology. I then turn to the question of feature accessibility in light of the debate between Representational Deficit Hypotheses and the Full Access approaches and their assumptions about L1 effects. In the same section, I also discuss the acquisition of individual morphemes and in Section 10.6, I review potential explanations for the periphrastic structure '*is + verb(x)*'. In Section 10.7, I discuss results on vocabulary and grammar altogether. Finally, I consider some methodological points regarding children's performance on the two task modalities in Section 10.8 before wrapping up in section 10.9.

10.1 Proficiency

It is important to evaluate whether the Russian and Chinese children in the four groups have matching proficiency, as expected by our initial design. Children's proficiency in English was assessed in a number of ways; first, through the Renfrew Word Finding Vocabulary task. Second, I considered the CEFR level of classes children attended in EF. Third, I used the narratives length (number of clauses), lexical diversity (type-token ratio), and syntactic complexity (number of dependent clauses) as additional measures to gauge proficiency.

These measures did not all discriminate between groups in the same way. The CEFR level of the classes children attended at the time of testing shows that close to half Russian 9-year-olds (14/32) are in an A1 CEFR class level. By contrast, just around an 8th of Chinese 9-year-olds (5/39) were at an A1 CEFR class level. Thus, according to the EF class teacher assessment and class placement, the number of A1 Russian children is 3 times the number of Chinese children and CEFR A1 (14 vs. 5).

The narrative length measure confirmed this picture, showing that Russian 9-year-olds produced significantly shorter stories than Chinese 9-year-olds. By contrast, the Renfrew task, the lexical diversity and syntactic complexity measures did not capture any difference between Russian and Chinese 9-year-olds.

This is perhaps due to the closer linguistic distance between English and Russian compared to English and Chinese[33] which might facilitate vocabulary and syntactic complexity for the Russian children, despite their lower overall proficiency.

The lower proficiency of Russian 9-year-olds is reflected in their accuracy with finiteness and confirmed by the statistical analyses in the previous chapter. Thus, the Renfrew score could not predict accuracy for the younger Russians, while it did for the younger Chinese children. This contrasts with narrative length, which appeared as a consistent predictor of accuracy across groups and L1s, indicating that appears a more discriminatory measure of L2 children's proficiency, at least for our groups.

There are two potential factors underlying the proficiency difference in the younger groups: the academic hours completed within 5 years in each country and more general aspects of education in the two countries. As seen in section 5.2.5, Russian 9-year-olds had completed on average fewer hours (around 25% less than all the Chinese counterparts; 592 vs 880). Further, in Shanghai, children start learning English in their day schools from the age 6 onwards attending English classes for 4 academic hours per week throughout primary school to 6 per week in secondary. By contrast, in Moscow, children start learning English at their day schools from the age 8 and for 3 hours per week also increasing in secondary school. Thus, Russian children start learning English later than Chinese learners and for fewer hours at their day

[33] Russian is an Indo-european language as is English, while Chinese pertains to the Sino-tibetan family of languages.

schools. Beyond English, literacy seems to start earlier for Chinese children than Russian children. These factors mean overall lower input for Russian children and later schooling which might explain their lower proficiency. It is worth noting that these differences do not impact the older groups which show matching proficiency as discussed below.

The CEFR level of the classes 12-year-old children attended at the time of testing was fairly similar; 31/34 Chinese children attended a B1 CEFR level class and 3/34 attended a B2 level class. As for Russians 39/42 attended a B1 CEFR level class and 3/42 attended a B2 level class. This shows that there is likely not too much of a difference in proficiency. Thus, the two groups are matched for proficiency, according to EF teachers assessment/class placement.

As with the younger groups, the teacher evaluations were confirmed by the narrative length measure which did not reveal a difference between the two groups. The vocabulary and syntactic measures, however, namely the Renfrew, lexical diversity and syntactic complexity yielded higher scores for the Russians, suggesting a higher proficiency. So again, it seems that these measures give Russian children an ‘advantage’ which, as speculated earlier, is potentially due to the linguistic and cultural proximity between Russian and English (I discuss further the notion of cultural proximity in the next section). Statistical analyses examining whether lexical diversity and syntactic complexity were predicted by Russian 12-year-olds’ higher vocabulary scores showed that this was not the case strengthening the explanation of linguistic proximity. I take the teacher evaluations and narrative length as more accurate estimations of children’s (overall) proficiency for the purposes of our study.

Considering the input of older children, Chinese and Russian 12-year-olds have similar average EF academic hours (CH:821 vs RU: 804). However, Chinese children may have had more input at their day school as they start English earlier and for more hours. In any case, these differences in input have not resulted in higher proficiency. The higher vocabulary and syntactic complexity scores of the Russians can be attributed to the linguistic proximity of English and Russian.

Following this discussion, we saw that narrative length was a more discriminatory measure of learners’ proficiency than the vocabulary task. It seems to be the case then that learners with L1s closer to the L2 may be favoured in vocabulary whereas this does not seem to hold for narrative length which appeared as a more objective measure of L2 proficiency not affected by linguistic distance. In fact, previous studies have shown that the text length distinguishes learners for writing proficiency (Ji-Young, 2014), length-related complexity measures (i.e. number of T-units and the number of clauses per T-unit) are good predictors of learners’ oral proficiency (Iwashita, 2006) and that more clauses are produced by higher proficiency young learners than less proficient ones (Hsieh & Wang, 2019). Thus, it seems that narrative length is a good predictor of L2 proficiency and stands as an objective measure. Vocabulary tasks like the present one though may provide distorted results in case we compare learners from different L1 backgrounds and differing considerably in linguistic distance from the L2.

10.1.1 Methodological issues; what does it mean to compare L2 children across countries

Following the discussion above, there are certain methodological issues worth pointing out with respect to the assessment of proficiency and the role of extralinguistic factors affecting children's rate of learning as emerged in the present cross-national study.

While the design of this study aimed to test children of exactly the same proficiency levels having set strict criteria (i.e. age of onset of learners learning English at EF, their age at time of testing keeping constant their length of exposure, same curriculum, same organisation: EF), we saw this was not exactly the case. Moreover, considering a number of different proficiency measures gave a mixed picture indicating that proficiency is a complex construct, not easily captured by one measure. In particular, vocabulary and syntactic complexity might be influenced by linguistic proximity.

To assess children's proficiency, it is then highly important to use a battery of measures to capture as many aspects of this multi-faceted construct as possible as individual measures might 'advantage' some groups over others due to linguistic proximity as is the case in the present study. To exemplify that, imagine we compare Dutch and Japanese learners of English. At early stages of acquisition, Dutch learners may score very high in a vocabulary task when their proficiency would not be as high yet. In other words, linguistic proximity may lead to inflated scores in vocabulary tasks that do not reflect the learners' true proficiency. This may hold with syntactic complexity as well as with other measures.

Our study suggests that teacher evaluations can be a reliable and perhaps indispensable measure of proficiency, especially for comparisons across different settings and nationalities, as teachers potentially have a more rounded view of children's performance including their macro-skills (e.g. speaking, reading, listening).

Cross-national studies are much needed, however, there might be further issues to be considered such as various extralinguistic factors affecting children's learning and progress with the L2. Specifically, we need to consider the L1 literacy and schooling as countries differ in their policies of when literacy starts and perhaps also the intensity of schooling. And this is only one aspect of what may be different. In fact, there may be socio-political and cultural variables affecting children's learning. I observed, for instance, while testing children, that there were few Chinese children knowing the word '*clown'* which appears in the Renfrew Vocabulary task whereas most Russians knew it. This is probably an effect of cultural distance between the Chinese and English which leads to lower L2 vocabulary scores and making the acquisition of 'cultural' vocabulary more challenging for Chinese learners. Finally, there seems to be more English available in media in Russia than China, which means that even though older children receive less English teaching in their day schools they might be exposed to more English outside their schooling.

10.2 L2 development for different ages of onset of learning English in EF

We have seen that in both Russia and China, with the exception of Russian 9-year-olds (RU_9: 592), all the other three groups have very similar mean academic hours of attendance (CH_9: 880, CH_12: 821, RU_12: 804) in EF. However, younger children have fewer teaching hours within 5 years than older ones, because older children will have consistently English in their day school for the whole 5-year period, while 9-year-olds will have English at their day school for 1 or 3 years depending on the country. As for the teaching hours in EF, in China the same hours are provided regardless of the age while in Russia hours increase with age meaning that older ones attend more English classes per week. These quantitative differences are complemented by qualitative differences in the curriculum followed from 4 to 9 and the curriculum followed from 7 to 12. For the younger ages, the EF curriculum seems to primarily aim for the child to develop motivation and love for the language. Children starting at 3 will be in classes targeting an A0 CEFR level up to the age of 6. In terms of grammar, they will be exposed only to very basic structures. Crucially, forms like simple past are introduced at the age of 7/8, irrespective of age of onset, so in our study, both 9- and 12-year-olds will have been introduced to past tense at almost the same age, 7/8. In older ages, the curriculum is more intense meaning that more structures are introduced within a year and there is more explicit teaching.

Because of these quantitative and qualitative differences in the input of younger and older children, the question of an earlier age advantage cannot be answered conclusively in this study. Younger and older children differ in terms of quantity of input and they also follow different curricula which may hide an earlier age advantage if this exists. Specifically, younger children lose the potential advantage of an earlier start as they are exposed to some crucial grammatical phenomena at the same age as older starters.

To conclude, it is important to acknowledge that, despite our controlled design, it is not easy in the end to disentangle biological age from the quantity and quality of input (hours and content) as well as socio-educational differences between China and Russia (e.g. English starting age at day school, literacy in mother tongue, parental attitudes towards education). Nevertheless, as we shall see, it is still possible to identify some age effects.

10.3 Age effects in L2 development of children with different ages of onset

One of the questions I asked in this study was whether a later age of onset will affect negatively children's performance -if we hypothesise that there is an optimal period in early childhood which fades away gradually. However, as we already saw it is impossible to reliably evaluate in this study whether there is an earlier age advantage.

The comparison between younger and older children did reveal some age effects as our statistical analysis showed that age did predict accuracy with older learners of both L1 cohorts being significantly more accurate on verbal morphology in all tests than younger ones. There was a clear age effect; older children were more accurate and more proficient than younger ones.

Further, age effects between 9 and 12-year-olds were also manifested through the use of the periphrastic marking across the L1s. Older children use the periphrastic structure less, which is instead more prevalent in the younger children (with the exception of Russian 12-year-old children on TEGI 3SG-agreement).

Bare forms also decrease with age. Older children use bare forms less than younger ones except for the Chinese 12-year-olds in 3SG whose percentage as we already saw remained high. Overall then, older children's performance is less varied and closer to the target system even when accuracy remains low.

The fact that older children consistently outperform the younger ones seems to confirm the rate advantage of older children in language learning that we know from previous studies (e.g. Muñoz, 2006; Myles & Mitchell, 2012; Pfenninger 2014; Jaekel et al., 2017), that is, older learners appear as faster learners than younger ones.

It is worth noting that immersed children learning English in a naturalistic context also produce the periphrastic structure (e.g. Ionin & Wexler, 2002; Paradis et al. 2008). Given that this structure does not exist in the input, it must be the result of learner internal mechanisms, that is, part of the implicit acquisition process. The fact that this pattern is prevalent in our younger children may suggest that younger children learn more implicitly than older ones showing patterns observed in immersed contexts and not explained by their instruction input.

Again, note that teaching for younger ages does not involve any explicit grammar teaching or focus on form. In this respect, teaching practice in younger ages supports implicit acquisition (probably tuning into to the younger children's learning style).

It seems then that various variables come into play and determine older children's performance. As we have already mentioned, these children have attended classes in their day schools for more years and have more advanced L1 literacy skills. Through schooling they have developed learning strategies that help their learning. They have more mature cognitive skills which are known to help general learning but also explicit language learning (Muñoz, 2006, 2010; Pfenninger, 2014; Jaekel et al., 2017). Finally, older learners have more advanced language aptitude skills (e.g. language analytic ability); recent research has shown that language aptitude is dynamic in childhood being developed up to the age of 12 (Roehr-Brackin &Tellier, 2019). We would thus expect older learners to have reached their maximum in that ability compared to younger ones or at least that they have higher language analytic ability. Thus, discussing age effects in instructed settings of limited input, we need to consider all these factors that contribute to older children's performance. In other words, any potential age advantage for the younger children may be overridden by the cognitive maturity, higher literacy skills and learning strategies of the older children.

Apart from these factors (e.g. metalinguistic knowledge, explicit teaching and feedback, cognitive abilities) though, the improvement attested in older children's performance also suggests that their grammars are constrained; they drop the periphrastic marking and start using the morphemes more entailing that through continuous exposure to input they come a step closer to acquiring the features.

To sum up, younger children's performance was quite diverse and showed more strongly patterns attested by L2 immersed children suggesting that younger ones learn English more implicitly. Older children's performance is more target-like, less diverse, and they seem to have dropped the periphrastic marking. Older starters can exploit their higher cognitive and general learning abilities to language learning -factors reported in literature to affect L2 learning.

10.4 Performance as an effect of limited input

As we discussed above younger children seem to learn more implicitly than older ones. We need, however, to examine in more detail whether the performance documented in this study is the result of more implicit acquisition of the younger children or whether it is just the result of the limited input children were exposed to. If input is the main contributing factor, then we would expect that immersed children in naturalistic contexts acquire English finite morphology without problems.

However, Li (2012) found similar patterns by examining six 7- to 9-year-old Chinese children immersed in the US (see Chapter 4, section 4.3.1). Specifically, the children had resided in the US between 4 and 5 months when first tested and were followed for over a 7-month period. They all attended English-speaking elementary schools since arrival and all of the children had been learning English as EFL in their Chinese schools prior to their arrival to the US for two to four years. Data were collected through a picture-based elicitation task as well as through a general conversation with the experimenter and were evaluated with respect to the production of copula and auxiliary *be*, 3SG -*s*, and past tense. Results showed that copula *be* was consistently supplied by all children to a very high percentage (93%). Auxiliary *be* in progressive contexts was supplied at a lower rate, although Li does not provide an average percentage of use for auxiliary *be*, as it is sometimes used with bare verbs and inflected verbs other than the progressive participle (that is, periphrastic marking instances as found in the present study). Children had 16% correct suppliance of 3SG -*s* in obligatory contexts, 13% correctly inflected verb with -*ed* in the regular past condition, and 38% correct suppliance of irregular forms. Their main errors on lexical verbs constituted of omissions while tense and number errors were mainly attested in the case of copula and auxiliary *be* (e.g. use of '*is'* instead of '*was'* in past context or use of '*is'* instead of '*are'*). These results are fairly similar qualitatively to what we have seen so far examining EFL children's performance on verb morphology in the present study.

Similar patterns were also found for Russian learners in the study by Ionin & Wexler (2002), (see Chapter 4, section 4.3.1). Twenty children with mean age 8;4 (range: 3;9-13;10) and varying lengths of exposure in the US (less than a year to up to three years) were assessed on copula and auxiliary *be*, third person singular -*s*, and past tense -*ed*. Results of spontaneous production data showed some omissions of inflection on '*be'* forms (cop.:16%, aux.: 33%) (note the higher percentage of omissions on auxiliary *be)* and a very high number of omissions of 3SG-agreement and past tense inflection on main thematic verbs (78% and 58% respectively)- with the 3SG-agreement -*s* marker being omitted more often than past tense. A small number of tense/agreement errors were also recorded specifically: 9% inappropriate use on copula *be*, 7% on auxiliary *be*, and 5% inappropriate use of 3SG -*s*. The authors also attested the overgeneration of *be* being used with bare verbs in a variety of contexts as we found in this study.

Finally, considering immersed children with more exposure to English, two studies (Jia & Fuse, 2007; Paradis et al., 2016) have shown that Chinese learners even after 5 or 6 years of exposure had not all mastered the 3SG-agreement and past tense morphemes.

Based on the above studies, it seems that inflectional morphology is challenging for immersed learners as it is for the EFL children of this study.

The comparison with Chinese and Russian immersed children then leads us to the conclusion that the difference between immersed and EFL children is not qualitative. Rather, the challenge with inflectional morphology appears as a matter of degree, which is probably due to the huge difference in the quantity and quality of input between immersed and instructed young learners.

Despite the general influence of input, it is noticeable that input measured in hours of teaching over a period of years in general did not impact on accuracy, with one puzzling exception; 12-year-olds with more hours of exposure/classroom attendance did in fact worse in main thematic verbs than those with fewer hours[34].

All in all, it seems that the children's performance -qualitatively speaking- is not an effect of input which would account for quantitative differences between the learners in the different contexts. The overall similarity between EFL and immersed children suggests that in both cases there are learner internal mechanisms in child L2 that are to a certain extent independent of the input and that, the difference between immersed and EFL children is primarily quantitative rather than qualitative. L2 children in both contexts seem to go through similar developmental stages. This is not a surprising result in view of e.g. early work by Ellis (1989).

10.5 The acquisition of finiteness by L2 children

Let us now turn to our question whether children after the age of 4 will resemble adult L2 acquisition of inflectional morphology rather than L1 acquisition patterns, as predicted by Meisel's hypothesis (2009).

First, we saw that L2 children in this study did significantly better in copula *be* than auxiliary *be* and in turn than main thematic verbs. Furthermore, another asymmetry was observed between the 3SG-agreement and past tense features. For L1 learners, finiteness morphemes emerge almost simultaneously (e.g. Brown, 1973; de Villiers & de Villiers, 1973; Rice et al., 1995; Rice & Wexler, 1996; Rice et al., 1998; Zobl & Liceras, 1994) and develop

---

[34] See Section 9.5 for a suggested explanation or speculation of this finding.

in parallel regardless of whether they are affixal or suppletive (e.g. Meisel, 2011) which is not the case here.

Second, we attested overgeneralisation of the progressive structure in the TEGI task where a habitual reading and structure was required. Overgeneralisation of the progressive structure has been found to characterise L2 acquisition (e.g. Hawkins & Casillas, 2008) but not the L1 acquisition. What we see in the L2 children of our study then resembles a lot aL2 acquisition.

Third, both Chinese and Russian L2 learners of English use overt subjects and non-finite forms which shows a lack of developmental relation between use of finiteness and overt subjects. This pattern also describes adult L2 acquisition of finiteness. The opposite pattern describes the L1 acquisition of English verb morphology; when children are in the OI or RI stage apart from omitting inflections, they also drop subjects (e.g. Hyams, 1986; Wexler, 1990; Wexler, 1998).

Fourth, L2 children omit verb inflection while using nominative subjects. Although I did not quantify the correct assignment of nominative case of pronominal subjects, it was clear already from the data transcription that accusative case e.g. him/her was very rarely -if at all- used. Previous research examining this issue has also shown that nominative case assignment of the pronominal subjects is generally unproblematic for L2 learners who acquire it early on (e.g. Li, 2012; Haznedar, 2007). This again contrasts with the OI or RI stage of L1 acquisition during which children use accusative case pronouns when using non-finite forms.

Finally, the periphrastic structure has also been found in various studies of adult L2 speech (see Chapter 4, Section 4.2). It has also been attested in L1 acquisition, however, by only a few studies (see Chapter 4, section 4.1) and very small samples, hence, not providing solid evidence about its properties and the underlying cause of its use.

Based on all these facts, we can conclude that cL2 acquisition of inflectional morphology resembles aL2 acquisition and not cL1 acquisition lending support to Meisel's hypothesis. It is a different type of acquisition from L1 acquisition as far as inflectional morphology is concerned.

10.5.1 Feature accessibility as a function of age of onset and L1

Having shown that inflectional morphology is indeed a vulnerable domain for L2 children, the next question is whether features of tense and agreement are accessible to them or not. Let us briefly summarise how each hypothesis accounts for optionality in L2 acquisition and its predictions about learners' performance:

Full Access Approaches

- The *Missing Surface Inflection Hypothesis* suggests that L2 children have fully specified syntactic representations: omissions of inflections indicate difficulty with mapping of the syntactic representations to morphology at PF; such difficulty is due to processing demands and communication pressures. This implies that problems learners have are production only problems which are superficial in nature. According to this hypothesis, L2 learners will omit inflections but will make few commission errors. When mapping to morphological forms fails, learners use the 'default' or 'elsewhere' form which, in our case are non-finite or bare verb forms. As finite forms are highly specified, L2 children do not overgeneralise them. Thus, under this approach children's performance should be characterised by omissions and few -if any- commission errors. Further evidence for this position comes from examining syntactic-related properties namely the acquisition of subjects which indicates acquisition of the syntactic representations.
- The *Prosodic Transfer Hypothesis* makes the same assumptions as the MSIH but adds prosody as a factor influencing the optionality of L2 inflections. Where languages differ in their prosodic structures, then problems are to be expected during development. To exemplify this, in Mandarin no adjunction to the prosodic word is permitted while in English this is how agreement or past tense marking is established. This difference in prosodic structures between the L1-Mandarin and the L2-English will lead to production difficulties for learners, specifically omissions in L2 production of verbs. Like the MSIH, the PTH predicts few commission errors or error patterns of other types (e.g. periphrastic marking). In addition, suppletive forms may be easier to acquire than regular affixation under this approach; if the difficulty learners have is accessing the morpheme *-ed*, for example, we would not expect them to have the same problem in case of an irregular past form.

Representational Deficit Hypotheses

According to proposals under the Representational Deficit Hypothesis (e.g. Contextual Complexity Hypothesis, Underspecification of AspP Hypothesis, Interpretability Hypothesis), L2 learners' early representations are characterised by feature inaccessibility which is manifested not only in omissions but crucially in overgeneralisation errors (e.g. present progressive use in 3SG habitual contexts) and error patterns different from L1 children's.

Let us now see again the findings characterising children's performance before evaluating each fact against the relevant hypotheses:

- (very) low accuracy
- high numbers of omissions
- asymmetry in the acquisition of copula be > auxiliary be > main thematic verb
- asymmetry in the acquisition of 3SG > 3PL
- asymmetry in the acquisition of 3SG-agreement > past tense
- no significant difference between regular/irregular verbs
- not important accuracy differences between verbs being marked for past tense syllabically or not
- overgeneralisation of the progressive structure in non-progressive contexts
- use of the periphrastic structure '*is + verb(x)*'

First, we saw (very) low accuracy across affixal morphemes (3SG-agreement and past tense) in both of the tasks by both L1 and age groups as well as a high number of omissions. The very low accuracy in inflectional morphology suggests that the morphemes have not been fully acquired especially if we consider a threshold of 90% accuracy. This finding considering the stage learners are at (i.e. late beginners- early intermediate) is consistent with both Full Access to UG and Representational Deficit approaches since they all aim to account for the optionality in use of inflection. Nevertheless, it should be mentioned that it is not clear at all when full access approaches expect ending of omissions. In principle, full access approaches cannot be compatible with omissions across developmental stages as they support nativelike representations.

L2 children did better on free morphemes rather than affixal inflections which is in line with the bulk of research in the field (e.g. Ionin & Wexler, 2002; Paradis et al., 2008; Hawkins & Casillas, 2008). It could just be a problem L2 children have with affixal morphology then. This could indeed be the case if children had fully mastered copula and auxiliary *be*. However, we attested a number effect, that is, when considering 3SG contexts performance was very high while for 3PL contexts -although few in number- performance appeared to be much more challenging. This points to a more general difficulty with verb morphology and finiteness in the early stages of child L2 acquisition; not to a problem specific to the inflection of main thematic verbs. This fact that copula and auxiliary *be* presented some challenges in 3PL contexts shows that the problem cannot just be phonological in nature – as copula and auxiliary *be* are suppletive forms. Thus, this finding speaks against the Prosodic Transfer Hypothesis (PTH). Could this number effect though be explained by the 'default/elsewhere' forms proposed by the Missing Surface Inflection Hypothesis (MSIH)? If we assume following their rationale that finite forms (3SG and past tense) are highly specified and non-finite forms may be underspecified and hence may substitute the former, we may expect '*be'* or '*are'* to work as default forms for copula and auxiliary *be*, rather than '*is'*; that is, because '*is'* is highly specified as is '*am'*, but '*are'* may be not as it is used in all other cases (i.e. 2nd person singular and all persons in plural). This would entail that 3PL contexts are easier and we would expect '*is'* in place of '*are'* when we actually see the reverse.

Another account with respect to the asymmetry between *be* forms and affixal morphology is offered by Ionin & Wexler (2002) who suggested that learners first associate morphological agreement with overt movement to T (i.e. *be* forms). This means that they do not initially analyse *-s* and *-ed* as inflectional morphemes; these require time to master due to establishing long-distance agreement (as is affixal inflection) which is an English-specific rule. They thus assume that learners have access to universal rules governing morphological agreement and that is why they master *be* forms which are overtly realised. Yet, they need time to master affixal inflection which is a language-specific rule. However, as Paradis et al. (2008) note such a proposal does not seem to be compatible with the MSIH as it would imply some incompleteness in morphosyntactic competence (that is, overt movement being realised earlier than covert movement).

It therefore appears that the MSIH cannot account for the difference between the morphemes. Instead, this ordering of morphemes is acknowledged and is attempted to be explained by Representational Deficit hypotheses as we shall see (e.g. Hawkins & Casillas, 2008).

Furthermore, in many studies *be* forms are considered altogether. Our study revealed that accuracy on copula *be* is significantly higher than accuracy on auxiliary *be,* a finding also attested in previous research (e.g. Paradis et al., 2008). The question is why copula *be* would be easier than auxiliary *be* if the challenge for learners is mapping to the forms; both representations map to the same form so that the difference must be related either to the syntactic features involved in each case or some other factor (e.g. processing difficulty). It is not clear how to accommodate the contrast between copula and auxiliary *be* within the MSIH and PTH. A factor other than syntactic representations would have to be evoked, e.g. higher processing difficulty for the auxiliary, that could explain the contrast in acquisition. Though such an account is not impossible, it is not currently obvious how to accommodate this asymmetry. By contrast, the Representational Deficit approaches do acknowledge such an asymmetry and offer some potential explanations. Hawkins & Casillas (2008), for example, in the framework of their Contextual Complexity Hypothesis, attribute the difference between accuracy on copula and auxiliary *be* to the fact that auxiliary *be* is restructured to accommodate an extra node in its Vocabulary entry, the interpretable [+/-progressive] feature. This extra node in relation to the copula *be* is what affects its retrieval strength and decreases its accuracy. Crucially, the additional complexity is the result of a more complex representation. In a similar vein, Gavruseva within her Underspecification of AspP hypothesis (2008) attempts to explain why auxiliary *be* is omitted more than copula *be* leading to an asymmetry in their acquisition: copula *be* is free of syntactic aspectual specification being specified only for tense/agreement. Its acquisition is therefore independent of Asp. The auxiliary *be* is further specified for the aspectual feature [-bound]. With syntactic aspectual features assumed to be underspecified early in L2 acquisition, this could explain why auxiliary *be* is omitted more than copula *be*. Although these are some possible explanations, we need, however, to consider that copula *be* is also associated with a stative interpretation and therefore is not aspectually free but it bears an atelic reading. Furthermore, progressive seems to be found in activities more entailing that it should be available to children's grammar as a feature. Thus, another possible explanation is simply that copula *be* is a main verb while the other one is an auxiliary, that is, a functional element. This again presupposes a deficit in the syntactic representations.

In sum, it is hard to see how to explain the difference between copula and auxiliary *be* within Full access accounts since the same form is involved in both cases. By contrast, the representational deficit hypotheses offer more natural accounts as the contrast can be linked to the underlying syntactic representations.

The next finding was the lower accuracy with 3SG-agreement in comparison to past tense. Again, the MSIH does not offer an account for it although it is consistent with it. One way to account for this contrast within MSIH would be to suggest that for the present the 'elsewhere' case is the bare form in all but one condition. By contrast, in the past all persons are associated with one form, which might make the mapping to *-ed* forms easier for the learner. While this is a plausible account, note that there is no difference between regular and irregular past, suggesting that the difficulty with inflectional marking is not necessarily associated with morphological realisation. Within the Representational Deficit approaches, Hawkins & Casillas (2008) argue that the reason for this discrepancy in early L2 grammars is the number of sister nodes in their vocabulary entries with 3SG having more sister nodes than past tense. To explain this more, they suggest that vocabulary entries for L2 learners are not specified as bundles of features as for L1 speakers but as statements about the terminal nodes into which a form is inserted. As such, *-ed* is specified for tense while *-s* is further specified for 3rd person and singular number. Underlying Hawkins & Casillas' assumption, is the view that uninterpretable syntactic features are absent from early[35] syntactic L2 representations. Tsimpli & Dimitrakopoulou (2007) propose that uninterpretable features cause higher difficulty in acquisition if absent from L1, in comparison to interpretable features, which are taken to be universally available to learners and, therefore, easier to acquire. Thus, uninterpretable features such as 3SG-agreement are more challenging for learners in comparison to past tense which involves interpretable features. Both Hawkins & Casillas as well as Tsimpli & Dimitrakopoulou agree on that: uninterpretable features which are not activated early in life through the learners' L1 will show persistent difficulties due to maturational effects, hence the asymmetry observed.

A further finding was that regular and irregular verbs did not differ in accuracy. The PTH would perhaps expect an asymmetry in favour of irregular verbs because suppletive forms can be learnt on an individual basis and as other lexical items. The absence of asymmetry suggests that both types of forms follow the same trajectory. As for the MSIH, if the non-finite form of a verb is considered the default/elsewhere form of both regular and irregular verbs, then this finding is neutral. For Representational Deficit hypotheses, the problem with both regular and irregular verbs is explained by a deficient representation; both regular and suppletive forms are therefore expected to be similarly affected which is what we see.

It was also found that whether a regular verb was marked for past tense syllabically (e.g. painted) or not (e.g. kicked) did not play an important role in their accuracy. It was only one verb found to be particularly problematic which was marked for past tense syllabically and this was the verb *lift-lifted*. This was a difficult item for participants as many of them confused it with the verb *leave* providing the form *left* or /*lefted*/. This again speaks against a phonological account of L2 optionality as following the PTH we would expect differences between those verbs as their prosodic structure differs. As for the MSIH, this finding is neutral as the default form would be the bare. According to RDH, this finding is expected as they assume that it is the feature of past tense that may cause difficulties and not the forms per se.

Next, in the results from TEGI, we saw that L2 children overgenerated the progressive structure in 3SG habitual contexts. The overgeneralisation of progressive aspect is a finding charactering cL2 and aL2 acquisition. This overgeneralisation of progressive aspect in habitual contexts perhaps occurs because a single present tense verb form covers both as it does in many

---

[35] Hawkins & Casillas (2008) do not specify what exactly they take 'early' to mean. It seems that it refers to beginner levels and perhaps any level before any restructuring of L2 syntactic representations may take place.

languages (e.g. in Russian for present tense). It could also be explained as dissociation between forms and meanings in L2 development; e.g. a learner needs to acquire both the form of e.g. the progressive tense as well as its meaning which determines the contexts it can be used in. This commission error could indicate a mismatch between the syntactic features associated with the progressive, potentially indicating underspecification of aspect in the mental representations. With either explanation, this finding could not be straightforwardly accommodated by MSIH or PTH as these hypotheses generally assume correct syntactic representations. Instead, it can be more easily accounted for by RD hypotheses.

Finally, we attested the periphrastic marking '*is + verb(x)*'; this pattern, very productive with younger children, is challenging for all accounts as it goes beyond optionality between bare forms and target inflections, which is what most hypotheses aim to account for. I will come back to this issue later on discussing whether it is a finiteness marker as proposed by Ionin & Wexler (2002) or not and evaluate the MSIH and Representational Deficit hypotheses based on further evidence from this structure.

To summarise, some of our key findings, in particular the better performance on copula *be* than auxiliary *be* and the difference between 3SG-agreement and past tense indicate problems with the acquisition of the underlying syntactic features and could not be straightforwardly accommodated by full access hypotheses. At the same time, the MSIH is consistent with differences in rates in the acquisition of morphemes. Further, for a conclusive evaluation we would need to test comprehension, expecting that children will have no problems in comprehension tasks if their problem is not their underlying representations but the surface (PF) forms. Comprehension tasks would provide the forms. (see Chapter 4, section 4.3.1 for a discussion of a comprehension study by Ionin and Wexler).

Representational Deficit hypotheses appear better placed to account for the contrasts between copula and auxiliary *be* and 3SG-agreement and past tense which both indicate issues with the underlying syntactic representations. However, if features are inaccessible how can we explain that children sometimes do use inflection and that they overgenerate *be* forms? Would the overgeneration of *be* forms suggest that agreement/tense features are acquired before aspectual features? Does the optionality suggest optional features in the syntactic representations? We also need to ask whether features remain inaccessible, and if increasing proficiency and the learners' L1 might influence the outcomes of our tests.

Turning to L1 effects in the acquisition of finiteness comparing Russian and Chinese learners, let us consider the predictions made by Full Access approaches and Representational Deficit Hypotheses.

According to the Full Transfer/ Full Access hypothesis L2 learners will rely on their L1 at initial stages before restructuring to the target system given that features are accessible. Thus, Chinese may have more protracted development compared to Russians in acquiring the morphemes due to the absence of tense/agreement marking. Both groups are expected to ultimately acquire the features.

Representational Deficit approaches do not expect L1 effects early on (Hawkins & Casillas, 2008) although they do not exclude them either (Tsimpli & Dimitrakopoulou, 2007); L1 influence is to become apparent as a function of increasing L2 proficiency as learners may have to resort to their L1 if (uninterpretable) features cannot be activated (Hawkins & Casillas, 2008). The assumption is that only features existing in the learners' L1 can be restructured through continuous exposure to input this concerning only uninterpretable features – interpretable features are assumed to be available to all learners (Tsimpli & Dimitrakopoulou, 2007; Hawkins & Casillas, 2008). Russians are then expected to acquire the features that exist in their L1 earlier than Chinese who will show persistent difficulties due to the missing features in their L1.

Unfortunately, I cannot consider the younger learners in our discussion of L1 effects since, as we saw Russian and Chinese are not matched for proficiency rendering the discussion of L1 effects unreliable. The younger Chinese outperformed the younger Russians, but this higher accuracy of Chinese learners is probably a reflection of their higher proficiency.

Let us focus then on older children's performance.

- Accuracy

Individual results revealed a difference between L1 cohorts; 9 Russian 12-year-olds performed at ceiling (≥90%) inflecting the majority of the main thematic verbs they produced in narratives. By contrast, there was only one Chinese 12-year-old reaching 90% accuracy on the same task.

- Bare forms

Russians produced fewer bare forms than Chinese overall. In both TEGI tasks examining 3SG-agreement and past tense, Russians produced fewer instances of bare verbs (19.5% and 5% respectively) than Chinese children (42% and 8% respectively). The same holds for narratives; Russians used bare forms (34.5%) less often than Chinese (48%). This finding shows that as Russian is a highly inflecting language, children may be aware of or more sensitive to finiteness marking. Thus, although they experiment with various alternate structures to express tense-mainly the periphrastic one- they do not produce as many bare forms as Chinese learners. In Chinese though there is no overt tense/agreement marking and bare forms are the norm, which may be the reason behind their frequent occurrence in their L2 English. This shows an asymmetry between the performance of the two populations which likely results from L1 influence.

- Periphrastic structure

Then considering the '*is + verb(-x)*' structure, there are some differences between the two populations; in TEGI 3SG, Chinese children rely on both '*is + bare verb*' (e.g. '*A dentist is fix teeth'*) as well as the progressive structure '*is + verb*-ing' (e.g. '*The dancer is dancing'*). In the same task, Russians use more frequently the progressive structure. In TEGI past, Chinese rely more on '*is + verb*-ing' while Russians use mainly 'i*s + inflected for past verb*' (e.g. '*He is brushed his hair'* or '*She is climbed*.').

Finally, in exploring the periphrastic structure in narrative production we also saw that '*is + bare verb*' was the most common pattern across the groups, but it was used less by Russians who also use '*be + verb inflected for past'*. Thus, although differences between the two populations with respect to the periphrastic structure are limited, there may still be some L1 influence within this structure.

What do these L1 effects tell us? Do features remain absent or do they get activated as a function of increasing proficiency and despite of the learners' L1?

Russians use more inflected forms than Chinese suggesting they are more sensitive to inflection. Now if we further consider the fact that many more individual Russians do perform at ceiling on verb morphology especially on main thematic verbs in narratives compared to Chinese children, this may mean that features existing in the learners' L1 are activated with increasing proficiency, and those features are acquired by Russians. For Chinese children the picture is different. Recall that proficiency could not explain accuracy in TEGI 3SG-agreement for Chinese 12-year-olds while it was a significant predictor for Russian 12-year-olds. It seems that Chinese may have persistent problems especially with 3SG-agreement.

Based on all the facts above, it seems that there are some L1 effects -yet not very pronounced. Will Russians acquire both 3SG-agreement and past tense then as the features exist in their L1? Will Chinese acquire the features at all -which are missing from their L1 and have not been activated early in life-, learn them in some way or do they show signs of fossilisation?

Starting with Russians, it seems that features as they exist in their L1 are acquired. Considering individual results, there were many already performing at ceiling. However, 3SG-agreement may need more time and exposure until it is fully acquired. Russians do seem to have some problems with 3SG-agreement as they use it to a limited extent in both task modalities while they overgeneralise the progressive aspect in 3SG habitual contexts. Although the feature of 3SG-agreement does exist in their L1, the delay observed may have to do with the English irregular system of agreement marking per se. Another potential explanation is that the feature configuration as far as grammatical aspect is concerned is different between Russian and English and perhaps learners' problem lies there -in the reconfiguration- requiring more time than past tense. Finally, the fact that Russians do worse than Chinese in 3SG-agreement in TEGI task may be attributed to the impact of the teaching methods followed at their day schools (whether it is more grammar-focused or more communicative) and to the fact that Russians receive less instructed input due to their later age of onset of learning English at their schools.

As for Chinese children, the facts that: 1) vocabulary does not predict accuracy in TEGI 3SG-agreement, 2) bare forms do not drop considerably in the TEGI 3SG-agreement, 3) they seem to mainly rely on the bare form in the periphrastic structure, 4) the increase in accuracy between younger and older learners is less pronounced than it is for Russians, as well as the fact that 5) considering individual results almost none performed at ceiling on inflectional morphology in narrative production, all point to the direction that Chinese children may have persistent problems with these features that do not exist in their L1. This picture is familiar from immersed children. As we saw in Chapter 4, not all Chinese children reached ceiling performance on 3SG-agreement and past tense even after 5-6 years of learning English naturalistically and attending English schools (e.g. Jia & Fuse, 2007; Paradis et al., 2016). However, no L1 activation of features does not mean no learning. Learning is of course available no matter the age. The activation of features though seems to be time-restricted when these features do not exist in one's L1. In other words, features not activated through a learner's L1 may be subject to maturational effects. Explicit or inductive learning seems to be the route Chinese learners will follow or that they have already followed – many Chinese children performed at ceiling in TEGI tasks but not in freer production, narratives (I discuss the difference in task modalities in section 10.7). I will come back later on this point when discussing how individual morphemes are acquired.

Considering again the two hypotheses, it seems that the Full Access approaches although not strongly disconfirmed cannot account for the present data. Instead, the results can be better explained by Representational Deficit approaches which assume that functional categories are underspecified and features may be missing in the early stages of L2 development.

10.5.4 Morpheme effects

The final question I aimed to address in this study was how individual morphemes are acquired. Again, there are two views in this respect, one following the Full Access approaches, and another assuming Representational Deficits. Let us first see each in turn and discuss their tenets and predictions.

According to the MSIH, the specific phenomena under investigation: the 3SG-agreement and the past tense are supposed to be fully specified forms while a bare verb is a 'default/elsewhere' form to be inserted in all other cases. We need to note though a potentially important difference. For 3SG-agreement the 'elsewhere' condition is within the present tense paradigm while for past, the elsewhere bare form is not part of the past paradigm. As the non-finite verb is underspecified with respect to finiteness, it can substitute finite forms because they are easier to access especially under processing demands and communication pressures. Following this hypothesis, we would expect both morphemes (3SG *-s* and *-ed*) to be equally affected. While we may further expect irregular verbs in past tense to show higher percentages of accuracy because their retrieval is only dependent on the featural content of the hosting node which is supposed to be fully specified.

Tsimpli & Dimitrakopoulou (2007), on the other hand, drawing on the distinction between interpretable and uninterpretable features suggest that the uninterpretable features will be subject to age effects (maturational constraints) if not activated in the learners' L1 while the interpretable features are acquirable. This implies an asymmetry in the acquisition of features with 3SG-agreement expected to be more difficult than past tense during development but also in ultimate attainment.

Statistical analyses showed that both Chinese and Russian children, older and younger, did significantly better with past tense than 3SG-agreement. This lends support to Tsimpli and Dimitrakopoulou (2007)'s hypothesis according to which interpretable features would be less challenging than uninterpretable ones. We further saw statistically that the regularity status of the verb did not predict accuracy in past tense; in other words, both regular and irregular verbs showed similar percentages of use.

This last point also answers to a possible question of whether the asymmetry attested may be the result of distributional (in-)consistency in marking of the features of person agreement and tense (agreement marked only in the 3rd person singular, past tense marked on all persons). Under the MSIH, it may be assumed that this asymmetry arises due to the fact that past tense is more systematic in PF, thus more frequent in input and thus with practice of its morphological mapping easier to acquire. However, in this case, we would expect irregular forms to be more accurate because they do not depend on a morphological mapping in the same way as regular forms which we did not see.

Further, if we consider the fact that past tense is introduced much later than 3SG-agreement we may have expected that children would do better in the former. On the other hand, it could also be the case that the past tense despite being introduced later is taught for a longer period of time because of the regular and irregular verbs and thus may be more frequent in the input children receive.

Based on the above discussion, the problem seems to lie within the features per se and not with the distributional (in-)consistency of marking, lending support to the Interpretability Hypothesis. However, we cannot exclude that frequency effects might also play a role.

10.6 Possible explanations for the periphrastic structure 'is + verb(x)' structure

I now turn to discuss in detail the periphrastic structure *'is + verb(x)'* in light of the relevant hypotheses; Full Access approaches, Representational Deficit Hypothesis, but also considering whether it could be aspectual marking.

10.6.1 L1 influence

The construction '*is + verb(x)'* was produced by both cohorts which come from two typologically different language backgrounds, that is, any L1 effect is very limited. In addition, we know from previous research that this structure has been recorded in production by children of various L1s; Russian (Ionin & Wexler, 2002), Spanish (García Mayo et al., 2005), Chinese (Yang and Huang, 2004; Li, 2012), and other L1s (Paradis et al. 2008). Thus, L1 transfer does not seem to be the explanatory factor for this pattern.

However, let us discuss potential L1 influence in some more depth before discarding it altogether. Russian lacks a copula '*be'* and also does not have an auxiliary '*be'*[36]. Therefore, it is unlikely that there would be direct transfer from the children's L1 Russian. Turning to Chinese (both Mandarin and Cantonese), it is a language that only marks aspect and has no tense or agreement. Excluding irrelevant aspectual markers such as '*-le, -guo, -zhe'* because they follow the verb and are not periphrastic, the only marker that might influence production is the imperfective (or ongoing or durative) marker '*zai'* which precedes the verb and may be

[36] There is only an equivalent to auxiliary *be* form in the compound future tense.

equivalent to '*is*'. An example of the use of the ongoingness aspect marker in Chinese taken from Klein et al. (2000) follows:

(1) Lisi zai chuan yi-jian qunzi.
    Lisi ZAI put-on one-CL skirt.
    'Lisi is putting on a skirt'

It could be hypothesised that children transfer this item from their L1 to English with the periphrastic structure being an attempt to mark progressive grammatical aspect. If this is the case, we would need to explain why children use a progressive item in TEGI where test prompts required a habitual reading or a past tense reading. This hypothesis seems implausible because children should be aware of the distinction between progressiveness, habituality and perfective as these features exist in their language.

Another structure which is used in Chinese that could be the source of crosslinguistic influence is the '*shi … de*' structure:

(2) Yayi shi xiubu yachi de (ren)
    dentist COP fix tooth REL person
    'A dentist (is a person who) fixes teeth'

In this case, learners could transfer the copula '*is*' while omitting the relativizer. Parentheses mean that 'person' is optional in the actual utterance. This structure is used when there is a generic interpretation only. This structure could form a possible source into the English periphrastic construction. Still, it cannot accommodate the use of periphrastic in past.

In sum, L1 influence is not a likely explanation for children's performance. Let us then leave this possibility aside and turn to the possibility that the periphrastic structure is attracted by lexical aspect.

10.6.2 Lexical aspect

We saw particular links between morphology and aspectual verb class. Specifically, the periphrastic structure is used mainly with achievements and accomplishments. This makes it unlikely that the periphrastic structure is an erroneous attempt to mark progressive aspect as in that case we would expect this type of marking with all predicate types except statives. If anything, our periphrastic seems to show up with verbs that in native speech are likely marked for past tense suggesting that the intended tense -if any- is more likely to be the past tense. Further considering the fact that children sometimes use '*be + inflected for past verb*' strengthens the above conclusion.

10.6.3 An explanation based on the framework of the Underspecification of AspP Hypothesis

Gavruseva (2008) found the periphrastic structure to occur mainly with statives and punctuals noticing that utterances with overgenerated *be* are finite. She took the finiteness of these uses as evidence that L2 children start figuring out aspectual distinctions relying on the lexical aspect semantics. Statives and punctuals – within this framework- are specified for the [+/-telic] feature which is checked in AspP. They are then expected to be finite. The overgeneration of *be* with these aspectual verb classes is linked to the emergence of a lexical-based aspectual system in L2 children's grammar and entails that syntactic T-chain licensing and Asp go hand-in-hand. However, the present data do not support Gavruseva's findings with respect to the verbs the periphrastic structure appears with. We saw that both Chinese and Russian children overgenerated *be* forms mainly with accomplishments and achievements. Accomplishments are non-punctuals, while statives appear only to a limited extent in a periphrastic structure in our data. Thus, the current data do not conform with Gavruseva's analysis and findings.

10.6.4 Finiteness marking

Ionin & Wexler (2002) suggest that the overgeneration of '*be*' is used to denote finiteness marking; overgenerated *be* forms act as substitutes of tense/agreement features. At the same time, they assume that T is fully specified for both tense and agreement, *be* forms (including the *be* overgeneration instances) providing evidence for this and that problems learners face are restricted to affixal inflection. The asymmetry found between forms of *be* and affixal inflection on thematic verbs is attributed to the different verb-raising possibilities; children first associate morphological realisation with *be* forms which raise overtly to T and they then master the affixal inflection on unraised thematic verbs, the former considered a UG rule, the latter an English-specific rule which requires more time.

As Paradis et al. (2008) comment, though, this proposal of the precocious acquisition of *be* as suggested by Ionin & Wexler may entail some incompleteness in morphosyntactic competence, this being against the main assumption of the MSIH which they adopt and perhaps more in line with representational deficit approaches.

In any case, this explanation in the frame of the MSIH still leaves a number of aspects of the current data unaccounted for.

In a similar vein, Parodi (2000, 2019) claims that although adult L2 learners have problems with inflection on thematic verbs, they rely on non-thematic verbs and dummy auxiliaries to mark finiteness, and that these are instances of T. She argues that L2 learners associate non-thematic verbs with grammatical information and thematic verbs with lexical information. The former being free morphemes are used as a strategy to mark finiteness when inflectional morphology has not been acquired yet. This strategy is further considered as economic within the minimalist theorising where ‘merge’ is seen as preferred to ‘move’. Data of the current study showed that children overgeneralised ‘*is*’ and no other light verbs. Further, we attested a gradual activation of features. Recall again, that 3SG contexts were much more accurate than 3PL contexts in both copula and auxiliary *be*.

It seems that the explanation that the overgeneration of ‘*is’* is a tense/agreement marker within a Full Representation account cannot accommodate the entirety of our data.

10.6.5 Towards a new hypothesis

Before actually discussing what the periphrastic structure denotes, let us briefly consider our findings from the previous chapter. We saw that L2 learners’ performance on verb morphology was characterised by 1) (very) low accuracy, 2) a number of asymmetries in order of emergence/acquisition: copula be > auxiliary be > main thematic verb inflection, past tense > 3SG-agreement, 3SG-contexts > 3PL-contexts. On the basis of these data, I argued that the features of tense and agreement are initially inaccessible while tense will be less challenging than 3SG-agreement because it is an interpretable feature and as such it will be acquired. As for 3SG-agreement which is uninterpretable I argued that this would perhaps depend on the learners’ L1.

Focusing on younger children’s performance of both L1 cohorts, we saw that at a stage when accuracy was very low, children either used bare forms or periphrastic marking instances. What does this say then for their underlying mental representations? What do children know and why do they overgenerate ‘*is*’ in these structures?

It seems that children have not associated the inflectional morphemes with syntactic features. Children have not associated the 3SG-*s* with the feature establishing agreement between the subject and the predicate. This is the reason especially younger children hardly use the 3SG-agreement and the past tense at all. In other words, children have not analysed the forms with -*s* and -*ed*. However, they start using the periphrastic structure ‘*is + verb’* with ‘*is’* being a finite element. My hypothesis then is that this ‘*is’* denotes the emergence of the finiteness marking category at a stage when the uninterpretable features are absent. Thus, although this seems a similar proposal to Ionin & Wexler’s or Parodi’s, it is based on different assumptions mainly on the absence of formal features early on and their potential future activation. The key difference then between the current hypothesis and Ionin & Wexler’s and

Parodi's hypothesis is the developmental aspect characterising cL2 acquisition of formal features speaking against their availability early on. In other words, the periphrastic marking is a first stage toward activation of formal features.

Children start noticing the verb to '*be'* as it is salient: it is very frequent (used both as a copula, as an auxiliary and as a dummy auxiliary), free, and always inflected. They come first to realise also through the progressive structure that '*is*' marks main thematic verbs for tense in English. While 3SG on e.g. she play-*s* means nothing to them, they start noticing the *be* form '*is*' as in e.g. she *is* playing. Children probably associate this '*is*' of the progressive structure with tense marking which is an interpretable feature. '*Is*' may thus work as a semantic trigger for the category finiteness or in other words, '*is'* is a placeholder of emerging finiteness at a stage when uninterpretable features are still absent. Thus, children may have acquired/activated the functional category of finiteness but not have acquired the paradigm yet in English. '*Is'* may therefore initially mark the syntactic position T but it does not mark a specific tense. This may explain why children sometimes use '*is + inflected for past verb*' in a past context.

This proposal entails that L2 acquisition proceeds differently from L1 acquisition; learners will first acquire the functional category through interpretable features and uninterpretable features are initially inaccessible as we have shown in previous sections; whether they may be activated depends on the learners' L1. Recall the various asymmetries attested when considering children's performance on the various morphemes; copula *be* > auxiliary *be* > main thematic verb inflection, past tense > 3SG-agreement, 3SG-contexts > 3PL-contexts. All these asymmetries show the gradualness characterising L2 acquisition. I thus speculate that in terms of syntactic structure while there may be a stage without any functional elements when children will not use any verb morphology at all, they will then activate the finiteness category manifested through the use of '*is + verb(x)*' before they acquire the paradigm -if they do, depending on their L1. Thus, the acquisition of L2 verb morphology and the syntactic representations do also seem to develop gradually from lack of features -> to interpretable features working as triggers of categories -> to uninterpretable features which may be activated as a function of increasing proficiency and the learner's L1.

The explanation holds for both populations regardless of their L1. However, it does not exclude the possibility of different notions to be encoded due to the L1. As such we saw Russians to use more the periphrastic structure with an inflected for past verb compared to Chinese who mainly used it with bare verbs. Although we did not test this statistically, it seems that Russians experiment more with it as they are more sensitive to inflection because of their L1. This particular instance may be a form-meaning mapping influenced by their L1 where they use a short adjectival type to form the past tense inflected for person, and gender. Chinese do not use as much inflection in this structure, again probably because it is also absent in their L1.

Overall, this explanation is better accommodated within Representational Deficit Hypotheses and not within Full Access approaches whose proponents would argue that features are all available from the beginning on.

10.7 Vocabulary and grammar learning

Let us now consider together findings on learners' performance on vocabulary and grammar. Results on Renfrew task as well as on TEGI tasks revealed some interesting patterns. As we saw, Russian children produced fewer correct 3SG-agreement instances (9y.o.: 2%, 12y.o.: 27%) compared to Chinese (9y.o.: 10%, 12y.o.: 42%) and '*be + verb-ing'* was a frequent error pattern in both groups; in Chinese, it is mostly found in the younger group but in Russian, both groups have it at over 30%. At the same time, in vocabulary (as shown in Renfrew task) the younger groups do not differ while the older groups differ with the Russians outperforming the Chinese. Thus, whereas older Russians perform lower in morphology of 3SG-agreement, they do better in vocabulary having significantly higher scores than Chinese.

These results may be due to when children start learning English at their day schools, the teaching method followed as well as the interplay between the two. First, if the teaching method were constant between the day schools of the two countries, then Russians who start later receive less instruction overall compared to Chinese who start earlier. This could have an impact on their low scores on 3SG-agreement compared to Chinese but on the other hand, it could have a positive impact on their learning of vocabulary because they started when they were 'readier' having higher memory skills and learning strategies. Second, the teaching method followed at school may be an important variable explaining these results if -apart from when children start learning English at their day schools- their learning is based more on explicit teaching or communicative approaches to language teaching. Unfortunately, we do not have such information to evaluate this potential explanation for children's performance but we need to acknowledge that this could have played a role to our findings. If the focus of the teaching method were more communicative where emphasis is given on "communicative competence" (Hymes, 1972; Liu, 2015) and meaning making rather than grammatical competence, then it may not be surprising why Russians do well in vocabulary and less well in grammar compared to Chinese counterparts. Similarly, if in Chinese schools, accuracy and focus on form is the focus of teachers, then again these results are viewed under new light. This is very speculative, of course, as we lack relevant information but we cannot ignore that the amount of instructed learning and the teaching method could be factors explaining what appears as a trade-off between learning vocabulary over learning morphology/grammar.

10.8 Test modalities and differences in children's performance

Before concluding the discussion, one final comment concerns the two test modalities used in this study; an elicited production task and a freer type of elicitation, a picture-based narration. Although children's performance was similar in many respects in the two modalities, in the sense that (very) low performance on finiteness marking was attested in both tasks, including large numbers of omissions, variability in performance and in errors, and use of the periphrastic structure '*is + verb(x)*', there were also some interesting findings and differences which are worth mentioning and discussing.

As far as accuracy is concerned, this appears overall lower in narratives than in the TEGI tasks. Further, considering individual results, while a considerable number of Chinese children performed at ceiling in the TEGI tasks, this was not the case in narrative production. There was only one Chinese child who reached 90% accuracy in terms of inflectional morphology measured as overt finiteness marking in the narrative production. The reason we see this difference in performance may be due to the nature of the two tasks; narratives are a more cognitively demanding and complex task as children have to compute many things together at the same time; such as vocabulary, structure, grammar, text coherence. Form may take less attention as children have to focus on all these aspects. In this sense, narratives may provide a more reliable picture of L2 children's knowledge. Elicited production tasks such as TEGI are more form focused and may make children think about the structure required and once a child realises that e.g. past tense is needed, s/he produces it in a consistent fashion almost automatically. Thus, it may be the case that the two tasks tap onto different types of learning; implicit vs explicit respectively. The task modality then may explain the difference observed in children's performance in the two tasks as well as the individual performance especially concerning older children.

Having said that, I do not want by any means to diminish TEGI's contribution to this study and research methodology in general. TEGI provided obligatory contexts and made it easier to interpret the overgeneralisation of progressive aspect in non-progressive contexts as well as the use of the periphrastic structure '*is + verb(x)*' with various meanings. In a narrative production task like the one used in the present study[37] this would be almost impossible as it is difficult to determine the temporal context the child has in mind and the tense that therefore is aiming to express. Obligatory contexts are important in research methodology as they may show whether form and meaning are dissociated which in narratives may be difficult to clarify.

Overall, both task modalities were equally important for research methodology. This finding comes to be added to numerous other studies (e.g. Paradis, 2010; Domínguez et al., 2012; López Prego & Gabriele, 2014; Tracy-Ventura & Myles, 2015) showing that task type affects learners' performance.

10.9 Summary

In this chapter, I discussed the findings from oral production tasks regarding the acquisition of finiteness in English by Chinese and Russian learners.

I first explained why the younger groups were not matched for proficiency despite our strict design. This discussion suggested that data fieldwork across two countries needs to take a broad view of assessment of proficiency incorporating extralinguistic factors that affect children's rate of learning. Vocabulary and syntactic complexity measures might be particularly sensitive to linguistic and cultural proximity while teacher evaluations emerged as reliable and indispensable source of evaluating proficiency.

I then considered in detail the quantity and quality of input children at the different age groups are exposed to and concluded that younger ones are exposed to less input in terms of both quantity and content. As a result, an earlier age advantage, if it exists, could be overridden by the richer input older children receive.

However, age effects were attested; it was found that older children's learning rate is faster -in line with previous research in the field. There was also some evidence that younger children learn more implicitly as they pattern more with immersed children while older ones showing a much less diverse performance on verb morphology appear to be better explicit learners and that may be the reason they do better overall; they can rely on other abilities and skills such as L1 literacy, cognitive and metalinguistic skills and they benefit more from teaching that focuses on form that younger children have not developed as much yet.

Age effects aside, all children's performance on inflectional morphology, specifically 3SG-agreement and past tense, was (quite) low and younger ones' more than older ones'.

[37] Of course, there are elicitation tasks which manipulate context to elicit specific structures.

However, the performance documented did not seem to be an effect of limited input as comparisons with immersed L2 children showed that they all go through similar developmental stages and their differences may be quantitative but not qualitative.

I then compared the data from our study to L1 acquisition and concluded that child L2 is more similar to adult L2 than L1, confirming Meisel's hypothesis according to which children older than 4 will pattern with L2 adults in the domain of inflectional morphology.

Considering further why L2 children resemble L2 adults and what this might entail for feature accessibility to tense and agreement, I argued that features of finiteness are inaccessible at the early stages of L2 development based on the following pieces of evidence: very low accuracy on inflectional morphology overall, a number effect manifested as challenging 3PL contexts in copula and auxiliary *be*, asymmetry between copula and auxiliary *be*, asymmetry between 3SG-agreement and past tense, overgeneralisation of the progressive tense in 3SG-habitual contexts. Based on all these facts, I argued that Full access approaches such as the MSIH and the PTH – although they are not disconfirmed- do not account sufficiently for these data which seem to be more consistent with representational deficit accounts.

Qualitative analysis of the production of older children showed only limited L1 effects. Their performance was in general terms similar, however, the fact that almost only Russian individuals performed at ceiling with respect to inflectional morphology in narratives and the use of less bare forms by Russians led me to hypothesise that perhaps Russians will follow different routes than Chinese learners; for Russians it may be a matter of time to activate the features existing in their L1 as they progress considerably with increasing proficiency while Chinese may have more persistent problems due to the missing features in their L1.

However, the asymmetry in the acquisition between 3SG-agreement and past tense features shows that they are the former i.e. the uninterpretable features (i.e. 3SG-agreement) that are harder to acquire than interpretable ones (i.e. tense) which are ultimately acquirable (performance on past tense was much better than performance on 3SG-agreement) lending support to the Interpretability Hypothesis (Tsimpli & Dimitrakopoulou, 2007).

Assuming that uninterpretable features are absent in the early stages of L2 acquisition and allowing for the possibility to be activated if existing in the learners' L1, I speculated that the use of the periphrastic structure is a stage in L2 development of English verb morphology which denotes the emergence of the functional category of finiteness. Children rely on semantic cues, that is, on interpretable features which are assumed to be available, mainly on '*is*' due to its salience (i.e. unbound, frequent, inflected) to activate uninterpretable features. The use of periphrastic structure then is a first stage leading to the activation of uninterpretable features.

If this is along the right lines, it may explain why Russians use more the periphrastic structure with inflected for past verbs; it may be a form-meaning mapping. L1 effects although not very pronounced are also encoded within this structure.

Finally, I considered task modalities concluding that different modalities are much needed in research as they can reveal us different aspects of children's L2 ability.

In the next chapter, I present the conclusions of this thesis and discuss potential limitations followed by suggestions for further research.

CHAPTER 11

CONCLUSIONS

11.0 Conclusions

This thesis aimed to contribute to a characterisation of child L2 acquisition by illuminating the impact of age of onset as well as the influence of the children's mother tongue on the acquisition of verb morphology in their L2 English. As already argued, this is an underresearched population which gained prominence during recent years. Focusing on EFL learners in a limited input context makes this research even more essential as literature on this field is scarce despite the fact that this population is interesting for both theoretical and practical reasons; theoretically, investigating EFL children allows us to evaluate the developmental stages these learners go through in their L2 and compare them with learners in immersed settings. This way, we can also explore the impact of input on the underlying processes of learning grammar. Further, EFL children constitutes a very large population of learners nowadays making it particularly important to further investigate as understanding their linguistic behaviour and the stages they go through can inform curriculum designers, teachers and stakeholders in general.

To address the research questions and hypotheses, Chinese and Russian learners aged 9 and 12 at time of testing with age of onset at 4 and 7 respectively were tested through elicited production tasks and narratives.

The first research question asked whether children older than 4 will pattern with L2 adults in the domain of inflectional morphology as Meisel (2009) has proposed. Considering the performance patterns of the L2 children and comparing them with L1 children and L2 adults, it was shown that L2 children older than 4 do not resemble L1 children and do not go through an OI/RI stage – in line with much research in the field. Instead, there were patterns such as the overgeneration of the progressive structure and the use of overt subjects with non-finite verbs which were similar to adult L2 learning lending support to Meisel's hypothesis.

We also showed in advance that their performance is not an effect of limited input. Comparing these children with naturalistic learners of English of similar ages and same L1s, we documented very similar performance and same error patterns. I suggested that the limited input can account for quantitative differences between EFL and immersed children but they all seem to go through the same developmental stages in the acquisition of inflectional morphology in English.

Addressing the question of whether age 4 is a cut-off point or rather is followed by a gradual decline in performance, younger and older starters were compared (recall the two groups had age of onset at 4 and were 9-year-olds at time of testing, the other two groups started learning English at 7 and were 12-year-olds at time of testing). No gradual decline in children's performance was attested; the reverse was actually revealed. Older children outperformed younger ones in all grammar tasks. The age effect was thus documented as a 'facilitating' factor for older children confirming their rate advantage. Focusing on younger and older children's performance and patterns, I argued that the two groups appear to learn somewhat differently; younger children resemble more immersed children meaning that they learn more implicitly while older children whose performance is less diverse and more restricted to the L2 target forms are assisted by a number of skills and abilities they have developed among which cognitive skills, learning strategies, language aptitude (or language analytic ability) being better explicit learners. The age factor then seems to be overridden by other abilities for these learners.

Note though that these results do not speak against an earlier age advantage. This could not be assessed reliably in this study as children do not follow exactly the same curricula while some phenomena are introduced much later for younger starters -at an age similar to that of older starters. Hence, any potential age advantage would vanish.

The rate advantage is also manifested to the overall proficiency levels of the age groups. Older ones were at a higher proficiency CEFR class level overall than younger ones. Further considering proficiency between L1 groups of the same age, this was found to be a quite complex construct and we needed a lot of measures to reliably assess it. It was found that younger learners are not in the same proficiency level but older starters are more comparable. This was particularly important when considering L1 effects as we shall see below. Before so let us proceed to the next issue of feature accessibility.

Having shown that inflectional morphology is problematic for these L2 children, the next question coming up was whether the problems children have are more superficial in nature as the Full Access approaches would argue or representational as the Partial Access or the Representational Deficit approaches suggest. In other words, the question was whether L2 children could access features of inflectional morphology, specifically features of tense and agreement. Based on children's performance and error patterns (i.e. very low accuracy, high numbers of omissions, asymmetry in the acquisition of copula be > auxiliary be > main thematic verb, asymmetry in the acquisition of 3SG > 3PL, asymmetry in the acquisition of 3SG-agreement > past tense, no significant difference between regular/irregular verbs, no effects of syllabic/non-syllabic past tense marking on accuracy, overgeneralisation of the progressive structure in non-progressive contexts, and use of the periphrastic structure '*is + verb(x)*') and interpreting each of them in light of the relevant hypotheses, it was shown that Full Access approaches although not strongly disconfirmed could not account for the entirety of these data. By contrast, the Representational Deficit approaches seem to better accommodate these results offering principled accounts for them.

To address the issue of feature accessibility then, it seems that features of inflectional morphology are initially absent. However, interpretable features will be first acquired than uninterpretable ones which are subject to L1 effects and maturational constraints. It was shown that all children did better in past tense which involves an interpretable feature than 3SG-agreement which involves uninterpretable features.

L1 effects were sought only for older learners as younger ones did not match for proficiency. Qualitative analysis of older Chinese and Russian children's performance (i.e. individual results, mean number of omissions, periphrastic structure) showed that while Russians can activate the features existing in their L1, this may not be the case for Chinese learners whose language lacks those features. Of course, we did not assess children's performance at their ultimate attainment but there were indications such as persistent bare form percentages, individual results revealing no ceiling performances, or the fact that proficiency could not explain 3SG-agreement in TEGI that led to that conclusion. Previous research showing that not all Chinese children who were immersed in English in the US for 5-6 years acquired the features of tense and agreement pointed to that direction as well. Examining Russians' performance, on the other hand, such as their lower number of bare forms, the ceiling performance of many individual learners on inflectional morphology in narratives, and the fact that they experiment more with inflection as shown in the periphrastic structure shows that features are acquired although their low percentages in 3SG-agreement may mean that a lot of exposure is needed before restructuring to the L2 target form.

Overall, child L2 acquisition of inflectional morphology seems to develop gradually. I speculated based on younger children's performance who use more extensively the periphrastic structure '*is + verb(x)*' that children go through two stages in the acquisition of finiteness features in English as revealed in the present data. There should be an earlier stage as well but we do not have enough data to support this. The first stage then is when finiteness emerges as a category while uninterpretable features are absent. Perhaps learners rely on semantic cues, that is, interpretable features to trigger this category. Children become aware of '*is*' because it is very salient, free, and always inflected. It is also attested before a thematic verb in the progressive structure which makes it even more visible and perhaps its role becomes more interpretable. It can accompany a thematic verb and is there to mark T. With increasing proficiency to the L2- that is at a higher proficiency level- children seem to drop the periphrastic structure and they move more towards the L2 target system. The overgeneration of *be* then signifies the stage toward the activation of uninterpretable features. Difficulties though may continue to exist especially for those phenomena not existing in the learners' L1 and hence not activated early in life.

11.1 Pedagogical implications of the study

This study aimed to investigate cL2 acquisition with mainly theoretical aspirations. However, it also aimed to provide the theoretical basis in order to inform stakeholders outside academia about the stages children go through when learning grammar in English in a limited input context. At this stage though we can only talk about pedagogical implications of this study leaving it to further work to offer more informed suggestions based on these results.

Considering the impact of hours as well as the way children learn English in different ages, we saw that academic hours did not significantly predict accuracy for younger children while there was a negative association for older children. Further, looking at learners' performance from a qualitative perspective we saw that younger children pattern more with immersed learners while older children appear to be better explicit learners. Putting all the pieces together and considering previous literature on the field showing that young children learn implicitly and that for implicit mechanisms to work a lot of hours are needed (e.g. Muñoz, 2010), it seems that more academic hours of classroom contact would be particularly beneficial for younger learners.

In addition, this could also facilitate the introduction of crucial grammar phenomena earlier for children who start at a very young age. Children immersed in English naturalistically get abundant input including even complex structures. There is no particular reason to assume that children of 4-5-years of age cannot learn the past tense, for instance, in an EFL context. It seems that EFL policy aims primarily to develop motivation and love for the language but perhaps young children would also benefit if exposed to more grammar early on.

With respect to the tasks used, we saw that while many Chinese children performed at ceiling at TEGI tasks, this performance was not replicated in the narratives. I argued that elicited production tasks perhaps draw learners' attention on certain structures making them aware of which one to use. Although this type of tasks may be good to raise awareness of a particular phenomenon and practice certain structures, they may tap more onto explicit learning. However, narratives are quite cognitively complex tasks which may tap more onto implicit learning. Narratives then could be an important instrument for teachers and language examiners to assess their learners' performance and progress. But apart from that, narrative as a task used in this study elicits more authentic discourse and is closer to real-world communication in the sense that it incorporates a communicative goal; telling a story. Teachers, curriculum/syllabus designers, and any other stakeholders could perhaps consider using oral narratives more in classroom practice if the ultimate goal of learning an L2 is -among others- to be able to communicate effectively in it, story-telling being one activity of communication and grammar its indispensable component. In fact, some scholars such as Robinson - and colleagues (Robinson, 1995; Robinson & Gilabert, 2007) - with his Cognition hypothesis suggests that increasing task complexity leads to higher accuracy. Jackson & Suethanapornkul (2013) in a meta-analysis of 9 studies on task complexity effects in learners' linguistic performance attest a small increase in accuracy. Many factors should of course be considered regarding the use of narratives such as the proficiency of the learners, their age, or other learner characteristics and narrative tasks need to be appropriately adjusted for particular learners. If targeting certain grammar structures the narrative tasks should be designed with that goal in mind. Tracy-Ventura & Myles (2015) showed that they managed to create more obligatory contexts for certain grammatical structures with narrative tasks designed for that purpose. To conclude, complex tasks such as narratives if designed carefully taking all factors discussed above into account and sequenced pedagogically -as suggested by the Cognition Hypothesis- may improve learners' accuracy in what can be considered a real-world communication activity.

11.2 Limitations of the present study & further research

The present study has hopefully contributed to a better understanding of cL2 acquisition in an EFL context addressing the issues of age effects and L1 influence by investigating in depth certain grammatical phenomena. There are some limitations though as well as fertile ground for further research.

First, the present study was a cross-sectional study so we could only get a glimpse of children's performance at a certain point in time. More longitudinal research in EFL learning of grammar could be very enlightening of children’s development illuminating the very initial stages as well as their ultimate attainment that we could not document in the present study.

Then, in this research, it was not possible to give a definite answer to the question whether earlier is better for grammatical acquisition as early and late starters follow different curricula and the introduction of phenomena took place when learners were at similar ages. It is necessary then further research to focus not only on the rate but also on the ultimate attainment of different starters in learning English focusing on grammar achievement controlling for curriculum effects and controlling for the timing in the introduction of the phenomena. Experimental designs can further enhance research findings in these real-life settings as they could fully control of variables that are not possible to do so in other cases.

Another limitation of the present study is the transcription of the data by the experimenter, that is, myself. Unfortunately, it was not feasible to have another transcriber/annotator that would allow us carry out an interrater reliability analysis. Although I used a professional audio-recording device and checked at least twice the transcriptions, it would be better research practice if another person would pass a sample through.

There is definitely a shortage of studies investigating grammar acquisition in EFL contexts focusing on micro-skills and not macro-skills. Further research with classroom learners needs to focus on grammar scrutinising specific grammatical properties.

More research in EFL contexts investigating the role of the L1 would be important to further clarify how the L1 impacts the L2 on more grammatical phenomena as this can also have practical consequences, allowing better preparation of material tailor-made for learners of different L1s and according to their needs.

Finally, further research may examine with more direct comparisons the interplay between age of onset and input. There are no many studies comparing ESL and EFL learners which will further need to establish that stages of development are the same for more grammatical phenomena and that input cannot change the route of development.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*, 481– 509.

Adger, D. (2003). *Core syntax: A Minimalist approach*. Oxford, England: Oxford University Press.

Adger, D. (2007). Three domains of finiteness: a minimalist perspective. In I. Nikolaeva (Ed.), *Finiteness: theoretical and empirical foundations* (pp. 23-58). Oxford, England: Oxford University Press.

Adger, D. & Svenonius P. (2011). Features in minimalist syntax. In C. Boeckx (Ed.), *The Oxford Handbook of Linguistic Minimalism* (pp. 27–51). Oxford, England: Oxford University Press.

Al-Hamad, M., Al-Malki, E. Casillas, G., Franceschina, F., Hawkins, R., Hawthorne, J., Karadzovska, D., Kato, K., Liszka, S.A., Lozano, C., Ojima, S., Okuwaki, N. & Thomas, E. (2002). Interpretation of English tense morphophonology by advanced L2 speakers. *EUROSLA Yearbook*, *2*(1), 49-69.

Andersen R. (1978). An implicational model for second language research. *Language Learning*, *28,* 221-82.

Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition, 16*, 133–156.

Andersen, S. W. & Shirai, Y. (1996). The Primacy of Aspect in First and Second Language Acquisition: The Pidgin-Creole Connection. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 527-570). San Diego, CA: Academic Press.

Armon-Lotem, S., Walters, J. & Gagarina, N. (2011). The Impact of Internal and External Factors on Linguistic Performance in the Home Language and in L2 among Russian-Hebrew and Russian-German Preschool Children'. *Linguistic Approaches to Bilingualism, 1*(3), 291–317.

Ayoun, D. (2013). *The Second Language Acquisition of French Tense, Aspect, Mood and Modality.* AILA Applied Linguistics Series 10. Amsterdam, NL: John Benjamins.

Ayoun, D. & Salaberry, R. (2008). The expression of temporality in English as a foreign language by French native speakers. *Language Learning, 58,* 555-595.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Bailyn, J.F. (2012). *The Syntax of Russian*. Cambridge, UK: Cambridge University Press.

Baker, M. (2008a). *The Syntax of Agreement and Concord.* Cambridge, UK: Cambridge University Press.

Bardovi-Harlig, K. (1992a). The relationship of form and meaning: A cross-sectional study of tense and aspect in the interlanguage of learners of English as a second language. *Applied Psycholinguistics, 13,* 253–278.

Bardovi-Harlig, K. (1995). A narrative perspective on the development of the tense/aspect system in second language acquisition. *Studies in Second Language Acquisition, 17,* 263–291.

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. (1999). From morpheme studies to temporal semantics. Tense-Aspect Research in SLA. *Studies in Second Language Acquisition, 21,* 341–382.

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use.* Oxford, UK: Blackwell.

Bardovi-Harlig, K., & Bergström, A. (1996). The acquisition of tense and aspect in SLA and FLL: A study of learner narratives in English (SL) and French (FL). *Canadian Modern Language Review, 52,* 308-330.

Bardovi-Harlig, K., & Reynolds, D. W. (1995). The Role of Lexical Aspect in the Acquisition of Tense and Aspect. TESOL Quarterly, 29(1), 107-131.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis, 20,* 3–49.

Blom, E., & Bosma, E. (2016). The sooner the better? An investigation into the role of age of onset and its relation with transfer and exposure in bilingual Frisian–Dutch children. *Journal of Child Language, 43*(3), 581–607.

Blom, E., & Paradis, J. (2016). Introduction: Special issue on age effects in child language acquisition. *Journal of Child Language, 43*(3), 473–478.

Bobaljik, J. D. (2017). Distributed Morphology. In *Oxford Research Encyclopedia of Linguistics*. Oxford, UK: Oxford University Press.

Bodrova, E. & Yudina, E. (2018). Early childhood education in the Russian Federation. In J. L. Roopnarine, J. E. Johnson, S. F. Quinn, & M. M. Patte (Eds.), *Handbook of* *international perspectives on early childhood education* (pp. 59-70). New York, NY: Routledge. https://doi.org/10.4324/9781315562193

Borer, H. (1984a). *Parametric syntax*. Dordrecht: Foris.

Brown, R. (1973). *A First Language: The Early Stages*. Mass: Harvard University Press.

Cabrelli Amaro, J., Campos-Dintrans, G., & Rothman, J. (2018). The role of L1 phonology in L2 morphological production. *Studies in Second Language Acquisition, 40*(3), 503–527.

Carroll, J. B., & Sapon, S. (2002). *Modern Language Aptitude Test: Manual.* Bethesda, MD: Second Language Testing.

Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code-mixing. In M. García Mayo, & M. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 77–93). Clevedon: Multilingual Matters.

Chen, S. Y., & Husband, E. M. (2018). Contradictory (forward) lifetime effects and the non-future tense in Mandarin Chinese. In *Proceedings of the Linguistic Society of America, 3*(1), 6.

Chen, X., & Meurers, D. (2016). CTAP: A Web-Based Tool Supporting Automatic Complexity Analysis. In *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity at COLING*, 113-119.

Chomsky, N. (1957). *Syntactic structures*. Mouton, ’s-Gravenhage.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981a). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding.* Cambridge, MA: MIT Press.

Chomsky, N. (1986a). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, N. (1995a). Bare phrase structure. In G. Webelhuth (Ed.), *Government and binding theory and the minimalist program* (pp. 383-439). Oxford: Blackwell.

Chomsky, N. (1995b). Categories and Transformations. In *Minimalist program*, Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The Minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step: Minimalist essays in honor of Howard Lasnik* (pp. 89–155). Cambridge, MA: MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 1–52). Cambridge, MA: MIT Press.

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and beyond: The Cartography of syntactic structures 3* (pp. 104–131). New York: Oxford University Press.

Chomsky, N. (2007). Approaching UG from below. In U. Sauerland & H.-M. Gartner (Eds.), *Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax- Semantics* (pp. 1–18). Berlin: Mouton de Gruyter.

Chomsky, N. (2008). On phases. In R. Freidin, C. P. Otero, & M. L. Zubizarreta (Eds.), *Foundational issues in linguistic theory*. *Essays in honor of Jean-Roger Vergnaud* (pp. 133–166). Cambridge, Ma: MIT Press.

Chondrogianni, V. (2018). Child L2 acquisition. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice, (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 103-126). John Benjamins Pub Co.

Chondrogianni, V., & Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Linguistic Approaches to Bilingualism, 1*(3), 318–345.

Clahsen, H., Eisenbeiss, S., & Vainikka, A. (1994). The seeds of structure. In T. Hoekstra & B.D. Schwartz (Eds.), *Language acquisition studies in generative grammar* (pp.85–118). Amsterdam: John Benjamins.

Comrie, B. (1976). *Aspect*. Cambridge, UK: Cambridge University Press.

Comrie, B. (1985). *Tense*. Cambridge, UK: Cambridge University Press.

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Cambridge, UK: Cambridge University Press.

Cubberley, P. (2002). *Russian. A Linguistic introduction*. Cambridge: Cambridge University Press.

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research, 28*(3), 369–382.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*(4), 499–533.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning, 55*, 1-25.

DeKeyser, R. M. (2013). Age Effects in Second Language Learning: Stepping Stones Toward Better Understanding: Age Effects in Second Language Learning. *Language Learning, 63*, 52–67.

Derrick, D. & Archambault, D. (2010). TreeForm: Explaining and exploring grammar through syntax trees. *Literary and Linguistic Computing, 25*(1), 53-66.

De Rosario-Martínez, H. (2015). *Phia: Post-hoc Interaction Analysis. R Package*. https://cran.r-project.org/web/packages/phia/phia.pdf

De Villiers, J. G., & de Villiers, P. A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research, 2*(3), 267–278.

Dewaele, J.-M., & Pavlenko, A. (2003). Productivity and lexical diversity in native and non-native speech: A study of cross-cultural effects. In V. Cook, (Ed.), *Effects of the second language on the first* (pp. 120-141). Multilingual Matters LTD.

Domínguez, L., Tracy-Ventura, N., Arche, M.J., Mitchell, R., & Myles, F. (2012). The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology. *Bilingualism: Language and Cognition*, *16*(3), 558-577.

Dowty, D. (1986). Tense and Aspect in Discourse. *Linguistics and Philosophy, 9*, 1-116.

Dugast, D. (1979). *Vocabulaire et stylistique. I Theatre et dialogue* [Vocabulary and style. Vol. 1 Theatre and dialogue]. Geneva, Switzerland: Slatkine-Champion.

Dye, C., Foley, C., Blume, M., & Lust, B. (2004). *The acquisition of verb morphosyntax: Evidence for the grammatical mapping paradigm.* Paper presented at the 78th Annual Meeting of the Linguistic Society of America, Boston, MA.

Dye, C., Foley, C., Blume, M., & Lust, B. (2004). Mismatches between Morphology and Syntax in First Language Acquisition Suggest a 'Syntax-First' Model. In *Proceedings of the 28th annual Boston University Conference on Language Development* (BUCLD 28), Boston, MA.

Eckart de Castilho, R., Mújdricza-Maydt, É., Yimam, S.M., Hartmann, S., Gurevych, I., Frank, A. & Biemann, C. (2016). A Web-based Tool for the Integrated Annotation of Semantic and Syntactic Structures. In *Proceedings of the LT4DH workshop at COLING 2016,* Osaka, Japan.

Eide, K. M. (Ed.) (2016). Finiteness and inflection: the syntax your morphology can afford. In *Finiteness Matters: On finiteness-related phenomena in natural languages* (pp.121–168). Amsterdam: John Benjamins.

Ellis, R. (1984). Can syntax be taught? A study of the effects of formal instruction on the acquisition of WH questions by children. *Applied Linguistics, 5*, 138-55.

Ellis, R. (1989). Are classroom and naturalistic acquisition the same? A Study of the Classroom Acquisition of German Word Order Rules. *Studies in Second Language Acquisition, 11*, 305-328.

Eubank, L. (1993/1994). On the transfer of parametric values in L2 development. *Language Acquisition, 3*, 183-208.

Gagarina, N., D. Klop, S. Kunnari, K. Tantele, T. Välimaa, I. Balciuniene, U. Bohnacker, & J. Walters (2012). *MAIN: Multilingual Assessment Instrument for Narratives*. ZAS Papers in Linguistics, 56.

García Mayo, M., & Garcia Lecumberri, M. L. (Eds.). (2003). *Age and the acquisition of English as a foreign language*. Multilingual Matters.

García Mayo, M., Lazaro Ibarrola, A., & Liceras, J. M. (2005). Placeholders in the English Interlanguage of Bilingual (Basque/Spanish) Children. *Language Learning, 55*(3), 445–489.

Gavruseva, E. (2002). Is there primacy of aspect in child L2 English? *Bilingualism: Language and Cognition, 5*(2), 109–30.

Gavruseva, E. (2003). Aktionsart, aspect and the acquisition of finiteness in early child grammar. *Linguistics, 41*, 723–55.

Gavruseva, E. (2004). Root infinitives in child second language English: an aspectual features account. *Second Language Research, 20,* 335–71.

Gavruseva, E. (2008). The development of copula and auxiliary be and overgeneration of be in child L2 English. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp.145-175). Amsterdam/Philadelphia: John Benjamins.

Gelderen, E. (2013). *Clause Structure (Key Topics in Syntax)*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139084628

Genesee, F., Paradis, J., & Crago, M. B. (2004). *Dual language development & disorders: A handbook on bilingualism & second language learning*. Baltimore, MD: Brookes.

Goad, H., White., L. & Steele, J. (2003). Missing inflection in L2 acquisition: defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics, 48,* 243-263.

Goad, H. and White, L. 2004: Ultimate attainment of L2 inflection: effects of L1 prosodic structure. In S. Foster-Cohen, M. Sharwood Smith, A. Sorace, A. & M. Ota (Eds.), *Eurosla Yearbook 4* (pp. 119-145). Amsterdam: Benjamins.

Goad, H., & White, L. (2006). Ultimate attainment in interlanguage grammars: a prosodic approach. *Second Language Research, 22*(3), 243–268.

Grano, T. (2017). Finiteness contrasts without Tense? A view from Mandarin Chinese. *Journal of East Asian Linguistics, 26*(3), 259–299.

Gries, S.T. (2009). Statistics for linguistics with R. A practical introduction. Berlin · New York: Mouton de Gruyter.

Grinstead, J. (2016). Root infinitives in child language and the structure of the clause. In J. L., Lidz, W., Snyder, & J., Pater, (Eds.), *The Oxford Handbook of Developmental Linguistics*. Oxford: Oxford University Press.

Guasti, M. T. (2016). *Language Acquisition: The growth of grammar*. MIT Press.

Flege, E. J. (2009). 'Give Input a Chance!'. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 175-190). Clevedon: Multilingual Matters.

Flynn, S. (1987). *A parameter-setting model of L2 acquisition*. Dordrecht: Reidel.

Filosofova, T., & Spöring, M. (2012). *Da! A practical guide to Russian grammar*. Oxon: Hodder Education.

Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S.J. Keyser (Eds.), *The view from building 20* (pp. 111–176). Cambridge, MA: MIT Press.

Harris, T. & Wexler, K. (1996). The optional-infinitive stage in child English. In H. Clahsen, (Ed.), *Generative perspectives in language acquisition* (pp. l-42). Philadelphia, PA: John Benjamins Publishing Company

Hartshorne, J.K., Tenenbaum, J.B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition, 177*, 263-277.

Hawkins, R. (2019). *How second languages are learned. An Introduction.* Cambridge, UK: Cambridge University Press.

Hawkins, R. & Chan, C. (1997). The partial availability of Universal Grammar in second language acquisition: the 'Failed Functional Features Hypothesis'. *Second Language Research, 13*(3), 187-226.

Hawkins, R., & S. Liszka. (2003). Locating the source of defective past tense marking in advanced L2 English speakers. In R. van Hout, A. Hulk, F. Kuiken, & R. Towell

(Eds.), *The lexicon-syntax interface in second language acquisition* (pp. 21-44). Amsterdam; Netherlands: John Benjamins.

Hawkins, R., & Casillas, G. (2008). Explaining frequency of verb morphology in early L2 speech. *Lingua, 118*(4), 595–612.

Haznedar, B. (1997). L2 Acquisition by a Turkish-speaking Child: Evidence for L1 influence. In E. Hughes & A. Greenhill (Eds.), *Proceedings of the 21st Boston University Conference on Language Development* (pp. 257-68). Somerville: MA: Cascadilla Press.

Haznedar, B. (2007). The acquisition of tense-aspect in child second language English. *Second Language Research, 23*(4), 383–417.

Haznedar, B. (2013). Child second language acquisition from a generative perspective. *Linguistic Approaches to Bilingualism, 3*(1), 26–47.

Haznedar, B. & Gavruseva, E. (2008). Recent perspectives in child second language acquisition. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp. 3-15). Amsterdam/Philadelphia: John Benjamins.

Haznedar, B. & Schwartz, B.D. (1997). Are there Optional Infinitives in child L2 acquisition? In E. Hughes, M. Hughes & A. Greenhill, (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development* (pp. 257-268). Somerville, MA: Cascadilla Press.

Housen, A. (2002). A corpus-based study of the L2-acquisition of the English verb system. In S. Granger, J. Hung, & S. Petch-Tyson, (Eds.), *Language Learning & Language Teaching 6* (pp. 77–116). John Benjamins Publishing Company.

Hsieh, C.-N., & Wang, Y. (2019). Speaking proficiency of young language students: A discourse-analytic study. *Language Testing, 36*(1), 27–50.

Hu, J., Pan, H., & Xu, L. (2001). Is there a finite vs. nonfinite distinction in Chinese? *Linguistics, 39*(6), 1117–1148.

Hu, B.Y. & Szente, J. (2009). Exploring the Quality of Early Childhood Education in China: Implications for Early Childhood Teacher Education. *Journal of Early Childhood Teacher Education, 30*(3), 247-262, DOI: 10.1080/10901020903084330

Huang, Y.Y. & Yang, S.Y. (1998). Acquisition of English tense-aspect markers by Chinese ESL learners. *Occasional Papers in Applied Language Studies* (Language Centre, Hong Kong Baptist University), III, 65-78.

Huang, Z. N. (2015). *On syntactic tense in Mandarin Chinese*. Paper presented at the 27th North American Conference on Chinese Linguistics (NACCL 27) (pp. 406–423). Los Angeles: UCLA.

Hulk, A., & Müller, N. (2000). Bilingual first language acquisition at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition, 3,* 227-244.

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.

Hymes, D. (1972). On communicative competence. In J.B. Pride & J. Holmes (Eds.), Sociolinguistics. Harmondworth: Penguin.

Hyltenstam, K. & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. J. Doughty, & M. H. Long, (Eds.), *The Handbook of Second Language Acquisition* (pp. 539-588). Oxford: Blackwell.

Ionin, T. (2008). Progressive aspect in child L2 English. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp. 17–53). Amsterdam: John Benjamins.

Ionin, T. & Wexler, K. (2002). Why is 'is' easier than '-s'? Acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research, 18,* 95–136.

Ionin, T., & Zyzik, E. (2014). Judgment and interpretation tasks in second language research. *Annual Review of Applied Linguistics, 34,* 37–64.

Iwashita, N. (2006). Syntactic Complexity Measures and Their Relation to Oral Proficiency in Japanese as a Foreign Language. *Language Assessment Quarterly, 3*(2), 151-169.

Jackson, D. O., & Suethanapornkul, S. (2013). The Cognition Hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning, 63*, 330–367.

Jaekel, N., Schurig, M., Florian, M., & Ritter, M. (2017). From early starters to late finishers? A longitudinal study of early foreign language learning in school. *Language Learning, 67*(3), 631–664.

Jarvis, S. & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York and London: Routledge.

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the united states. *Applied Psycholinguistics, 24*(1), 131–161.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech, Language, and Hearing Research, 50*(5), 1280–1299.

Ji-Young, K. (2014). Predicting L2 writing proficiency using linguistic complexity measures: A corpus-based study. *English Teaching, 69*(4), 27-51.

Johnson, J. & Newport, L. E. (1989). Critical Period Effects in Second Language Learning. The Influence of Maturational State on the Acquisition of English as a Second Language, *Cognitive Psychology, 21*, 60–99.

Klein, W. (2006). On Finiteness. In V. Geenhoven, (Ed.), *Semantics in Acquisition* (pp. 245-272). Springer, Dordrecht.

Klein, W., Li, P., & Hendriks, H. (2000). Aspect and assertion in Mandarin Chinese. *Natural Language & Linguistic Theory, 18*(4), 723–770.

Klima, E., & Bellugi, U. (1966). Syntactic regularities in the speech of children. In J. Lyons, & R. Wales, (Eds.), *Psycholinguistic papers* (pp. 183–208). Edinburgh: Edinburgh University Press.

Koeneman, O., & Zeijlstra, H. (2017). *Introducing Syntax*. Cambridge, UK: Cambridge University Press. doi:10.1017/9781316156391

Krashen, S., Long, M., & Scarcella, R. (1979). Age, Rate and Eventual Attainment in Second Language Acquisition. *TESOL Quarterly, 13*(4), 573-582. doi:10.2307/3586451

Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press.

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.

Kroffke, S. & Rothweiler, M. (2006). Variation im frühen Zweitspracherwerb des Deutschen durch Kinder mit türkischer Erstsprache [Variation in early second language acquisition of German by children with Turkish as L1]. In M. Vliegen, (Ed.), *Proceedings of the 39th Linguistics Colloquium* (pp. 145–153). Amsterdam: Peter Lang.

Lardiere, D. (1998a). Case and tense in the 'fossilized' steady state. *Second Language Research, 14*, 1–26.

Lardiere, D. (1998b). Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research, 14*, 359–75.

Lardiere, D. (1998c). Parameter-resetting in morphology: evidence from compounding. In M.-L. Beck, (Ed.), *Morphology and its interfaces in second language knowledge* (pp. 283–305). Amsterdam: John Benjamins.

Lardiere, D. (2000). Mapping features to forms in second language acquisition. In J. Archibald, (Ed.), *Second language acquisition and linguistic theory* (pp. 102-129). Malden, MA: Blackwell.

Lardiere, D. (2003). Second language knowledge of [±past] vs. [±finite]. In *Proceedings of the 6th generative approaches to second language acquisition conference* (GASLA) (pp.176-189). Somerville, MA: Cascadilla Proceedings Project.

Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case study.* Lawrence Erlbaum; Mahwah, NJ.

Lardiere, D. (2008). Feature-Assembly in Second Language Acquisition. In J. Liceras, H. Zobl, & H. Goodluck, (Eds.), *The role of formal features in second language acquisition* (pp. 106-140). New York: Routledge.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research, 25*(2), 173–227.

Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second Language Research, 24*, 35–63. doi:10.1177/0267658307082981

Lee, N., & Huang, Y. Y. (2004). To be or not to be - The variable use of the verb BE in the interlanguage of Hong Kong Chinese children. *RELC Journal, 35*(2), 211–228.

Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley and Sons.

Li, C. N. & Thompson, S. A. (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

Li, H. (2013). Teaching Chinese Literacy in the Early Years: A Comparison of L1 and L2 Preschool Classrooms in Shenzhen and Singapore. *Asia-Pacific Journal of Research in Early Childhood Education, 7*(3), 19-43.

Li, H., Corrie, L.F., & Wong, B.K.M. (2008). Early teaching of Chinese literacy skills and later literacy outcomes. *Early Child Development and Care, 178*(5), 441–459.

Li, M.-C. (2012). *The acquisition of tense and agreement by early child second language learners* (Doctoral dissertation). University of Illinois, Urbana-Champaign.

Li, S. (2015). The Associations Between Language Aptitude and Second Language Grammar Acquisition: A Meta-Analytic Review of Five Decades of Research. *Applied Linguistics, 36*(3), 385–408. doi:10.1093/applin/amu054

Liceras, J. M. (2010). Second Language Acquisition and Syntactic Theory in the 21st Century. *Annual Review of Applied Linguistics, 30,* 248–269.

Liceras, J., Zobl, H., & Goodluck, H. (Eds.). (2008). *The Role of formal features in second language acquisition*. New York: Routledge.

Lin, J. (2003). Temporal reference in Mandarin Chinese. *Journal of East Asian Linguistics, 12*(3), 259–311.

Lin, J.-W. (2006). Time in a language without tense: The case of Chinese. *Journal of Semantics, 23*(1), 1–53.

Lin, J. (2010). A tenseless analysis of Mandarin Chinese revisited: A response to Sybesma 2007. *Linguistic Inquiry, 41*(2), 305–329.

Lin, T. H. (2011). Finiteness of clauses and raising of arguments in Mandarin Chinese. *Syntax, 14*(1), 48–73.

Lin, T. H. (2012). Multiple-modal constructions in Mandarin Chinese and their finiteness properties. *Journal of Linguistics, 48*(1), 151–186.

Lin, T.-H. J. (2015). Tense in Mandarin Chinese sentences. *Syntax, 18*(3), 320–342.

Liu, S. (2015). Reflections on Communicative Language Teaching and its application in China. *Theory and Practice in Language Studies, 5*(5), 1047-1052.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12,* 251-285.

López-Prego, B. & Gabriele, A. (2014). Examining the impact of task demands on morphological variability in native and non-native Spanish. *Linguistic Approaches to Bilingualism, 4*(2), 192-221.

Mai, Z., & Yuan, B. (2016). Uneven reassembly of tense, telicity and discourse features in L2 acquisition of the Chinese shì … de cleft construction by adult English speakers. *Second Language Research, 32*(2), 247–276.

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In *UPenn Working Papers in Linguistics, 4*, 201–225.

Matthewson, L. (2006). Temporal semantics in a superficially tenseless language. *Linguistics and Philosophy, 29*(6), 673–713.

Meisel, J. M. (1991). Principles of Universal Grammar and strategies of language learning: Some similarities and differences between first and second language acquisition. In L. Eubank, (Ed.), *Point counterpoint: Universal Grammar in the second language* (pp. 231–276). Amsterdam: Benjamins.

Meisel, J. M. (2009). Second language acquisition in early childhood. *Zeitschrift für Sprachwissenschaft, 28*(1), 5-34. <doi:10.1515/ZFSW.2009.002>

Meisel, J. M. (2011). *First and second language acquisition: Parallels and differences*. Cambridge: Cambridge University Press.

Mezhevich, I. (2008). A feature-theoretic account of tense and aspect in Russian. *Natural Language & Linguistic Theory, 26*(2), 359–401.

Mitchell, R., Myles, F. & Marsden J. E. (2013). *Second language learning theories*, 3rd ed. Abingdon: Routledge.

Mobaraki, M., Vainikka, A., & Young-Scholten, M. (2008). The status of subjects in early child L2 English. In B. Haznedar & E. Gavruseva (Eds.), *Current trends in child second language acquisition: A generative perspective* (pp. 209–236). Amsterdam: John Benjamins.

Muñoz, C. (Ed.) (2006). *Age and the rate of foreign language learning: Second language acquisition.* Clevedon, UK: Multilingual Matters.

Muñoz, C. (2010). On how Age Affects Foreign Language Learning. In *Advances in Research on Language Acquisition and Teaching: Selected Papers,* GALA 2010.

Myles, F., & Mitchell, R. (2012). *Learning French from ages 5, 7, and 11: An investigation into starting ages, rates and routes of learning amongst early foreign language learners.* (ESRC End of Award Report, RES-062-23-1545). Swindon, England.

Nghia, K. (2006). *The Temporal Location in Chinese: Looking for the Temporal Adverb*. Paper presented at the Conference of the European Association of Chinese Linguistics (EACL4), Budapest.

Nikolaeva, I. (Ed), (2007). *Finiteness: theoretical and empirical foundations*. Oxford: Oxford University Press.

Nikolov, M., & Djigunovic, J. (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual Review of Applied Linguistics, 26*, 234–260.

Norris, J. M., & Ortega, L. (2009). Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied Linguistics*, *30*(4), 555–578.

Noyer, R. (1997). *Features, positions, and affixes in autonomous morphological structure*. New York: Garland.

Paradis, J. (2010). Bilingual Children's Acquisition of English Verb Morphology: Effects of Language Exposure, Structure Complexity, and Task Type: Bilingual Children's Acquisition of English. *Language Learning*, *60*(3), 651–680.

Paradis, J., Rice, M., Crago, M., & Richman, A. (2004). *Missing inflection or (extended) optional infinitives? Comparing child L2 English with English SLI*. Paper presented at the 29th Annual Boston University Conference on Language Development, Boston University.

Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics, 29*(4), 689–722.

Paradis, J. (2011). Individual differences in child second language acquisition. Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*(3), 213-37.

Paradis, J., Tulpar, Y., & Arppe, A. (2016). Chinese L1 children's English L2 verb morphology over time: individual variation in long-term outcomes. *Journal of Child Language, 43*(3), 553–580.

Parodi, T. (2000). Finiteness and verb placement in second language acquisition. *Second Language Research*, *16*(4), 355–381.

Parodi, T. (2019). Linguistic intuitions beyond grammatical correctness in untutored L2 learners. Paper presented at the MEITS Seminar, Cambridge, Cambridge University.

Penfield, W. & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Pfenninger, S. E. (2014). The misunderstood variable: Age effects as a function of type of instruction. *Studies in Second Language Learning and Teaching*, *3*, 529–556.

Pierce, L. J., Genesee, F., & Paradis, J. (2013). Acquisition of English grammatical morphology by internationally adopted children from China. *Journal of Child Language, 40*(5), 1076–1090.

Platzack, C. (1996). The initial hypothesis of syntax: a minimalist perspective on language acquisition and attrition. In H. Clahsen (Ed.), *Generative perspectives on language acquisition: empirical findings, theoretical considerations, crosslinguistic comparisons* (pp. 369–414). Amsterdam: John Benjamins.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Poeppel., D. & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language, 69*(I), l-33.

Pollock, J.-Y. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry, 20*, 365–424.

Prego, B. L., & Gabriele, A. (2014). Examining the impact of task demands on morphological variability in native and non-native Spanish. *Linguistic Approaches to Bilingualism 4*(2), 192–221.

Prévost, P. & White, L. (2000a). Accounting for morphological variation in second language acquisition: truncation or missing inflection? In M.A. Friedeman, & L. Rizzi, (Eds.), *The acquisition of syntax* (pp. 202-235). London: Longman.

Prévost, P. & White, L. (2000b). Missing surface inflection or impairment in second language? Evidence from Tense and Agreement. *Second Language Research, 16*(2), 103–133.

R Core Team (2014). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. http://www.R-pro- ject.org/.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax*. Oxford: Blackwell.

Radford, A. (2004). *English Syntax: An Introduction*. Cambridge: Cambridge University Press.

Radford, A. (2004). *Minimalist Syntax: Exploring the Structure of English.* Cambridge: Cambridge University Press. doi:10.1017/CBO9780511811319

Renfrew, C. (1998). *The Renfrew Language Scales: Word Finding Vocabulary Test*. Milton Keynes: Speechmark.

Rice, M. & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment. *Journal of Speech, Language and Hearing Research, 39*, 1236–1257.

Rice, M., and K. Wexler. 2001. *Test of Early Grammatical Impairment*. New York, NY: The Psychological Corporation.

Rice, M., Wexler, K. & Cleave, P. (1995). Specific Language Impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research, 38*, 850-863.

Rice, M, Wexler, K., & Hershberger S. (1998). Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1412–1431.

Rizzi, L. (1994). Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition, 3*, 371-393.

Rizzi, L. (1994b). Early null subjects and root null subjects. In T. Hoekstra, & B. Schwartz (Eds.), *Language acquisition studies in generative grammar* (pp.151-176). Amsterdam: Benjamins.

Rizzi, L. & Cinque, G. (2016). Functional Categories and Syntactic Theory. *Annual Review of Linguistics, 2*(1), 139-163.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Learning Language, 45*, 99–140.

Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *IRAL-International Review of Applied Linguistics in Language Teaching, 45*, 161–176.

Roehr-Brackin, K. & Tellier, A. (2019). The role of language-analytic ability in children's instructed second language learning. *Studies in Second Language Acquisition, 41,* 1111–1131. doi:10.1017/S0272263119000214

Roeper, T. & Rohrbacher, B. (1994). *True pro-drop in child English and the Principle of Economy of projection,* (Unpublished Manuscript), University of Massachusetts, Amherst.

Roeper, T. & Rohrbacher, B. (2000). Null Subjects in Early Child English and the Theory of Economy of Projection. In S.M. Powers, & C. Hamann, (Eds.), *The Acquisition of Scrambling and Cliticization* (pp. 345-396). Springer, Dordrecht.

Rothweiler, M. (2006). The acquisition of V_2 and subordinate clauses in early successive acquisition of German. In C. Lleó, (Ed.), *Interfaces in multilingualism: acquisition, representation and processing* (93–115). Amsterdam: John Benjamins.

Salaberry, R. & Shirai, Y. (2002). L2 acquisition of tense-aspect morphology. In M.R. Salaberry, & Y. Shirai, (Eds.), *The L2 Acquisition of Tense–Aspect Morphology* (pp. 1-20). John Benjamins Publishing Company.

Santelmann, L., Berk, S., and Lust, B. (2000). Assessing the Strong Continuity Hypothesis in the Development of English Inflection: Arguments for the Grammatical Mapping Paradigm. In *Proceedings of the West Coast Conference on Formal Linguistics 19* (pp. 439-452). Somerville: Cascadilla Press.

Santelmann, L., Berk, S., Austin, J., Somashekar, S., & Lust, B. (2002). Continuity and development in the acquisition of inversion in yes/no questions: Dissociating movement and inflection. *Journal of Child Language, 29*(4), 813–842.

Schubert, S. (2002). *'Are you wanting a cup of coffee?': Overuse of the progressive aspect in Indian English.* Konstazer Online-Publikations-System (KOPS).

Schütze C. & Sprouse J. (2013). Judgment data. In R. Podesva & D. Sharma, (Eds.), *Research methods in linguistics* (pp. 27-50). Cambridge, UK: Cambridge University Press.

Schwartz, B. D. (2004). Why child L2 acquisition? In *LOT Occasional Series 3* (pp. 47–66). Utrecht, The Netherlands: Netherlands Graduate School of Linguistics (LOT).

Schwartz, B.D. & Sprouse, R. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research, 12*, 40–72.

Shirai, Y. (1991). *Primacy of aspect in language acquisition: Simplified input and prototype* (Unpublished doctoral dissertation). University of California, Los Angeles.

Shirai, Y., & Andersen, R. W. (1995). The Acquisition of Tense-Aspect Morphology: A Prototype Account. *Language, 71*(4), 743.

Slabakova, R. (2016). *Second Language Acquisition*. Oxford: Oxford University Press.

Smith, C.S. (1983). A theory of aspectual choice. *Language, 59,* 479–501.

Smith, C.S. (1991). *The parameter of aspect*. Dordrecht: Kluwer.

Smith, C. S. (2008). Time with and without tense. In J. Gúeron & J. Lecarme (Eds.), *Time and modality* (pp. 227–249). Springer.

Sybesma, R. (2004). Exploring Cantonese tense. In L. Cornips, & J. Doetjes, (Eds.), *Linguistics in the Netherlands*. Amsterdam: John Benjamins.

Sybesma, R. (2007). Whether we tense-agree overtly or not. *Linguistic Inquiry, 38*(3), 580–587.

Sun, H., de Bot, K., & Steinkrauss, R. (2015). A multiple case study on the effects of temperamental traits in Chinese preschoolers learning English. *International Journal of Bilingualism, 19*(6), 703–725.

Sun, H., Steinkrauss, R., Tendeiro, J. & De Bot, K. (2016). Individual differences in very young children's English acquisition in China: Internal and external factors. *Bilingualism: Language and Cognition, 19*(3), 550–566. doi:10.1017/S1366728915000243

Tan, C. (2012). The culture of education policy making: Curriculum reform in Shanghai. *Critical Studies in Education, 53*(2), 153-167.

Tesan, G., & Thornton, R. (2004). Overspecification in early English. In J. van Kampen, & S. Baauw, (Eds.), *LOT Occasional Series 3* (pp. 469-480). Utrecht, The Netherlands: Netherlands Graduate School of Linguistics (LOT).

Thornton, R., & Crain, S. (2013). Parameters: the pluses and the minuses. In M. den Dikken, (Ed.), *The Cambridge handbook of generative syntax* (pp. 927-970). Cambridge: Cambridge University Press.

Tracy-Ventura, N., & Myles, F. (2015). The importance of task variability in the design of learner corpora for SLA research. *International Journal of Learner Corpus Research, 1*(1), 58-95.

Tsai, W.-T. (2006). *Tense Anchoring in Chinese*. Paper presented at the 4th Conference of the European Association of Chinese Linguistics, Budapest.

Tsimpli, I. M. (2014). Early, late or very late? Timing acquisition and bilingualism. *Linguistic Approaches to Bilingualism, 4*(3): 283-313. <doi: 10.1075/lab.4.3.01tsi>

Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The Interpretability Hypothesis: Evidence from Wh-Interrogatives in Second Language Acquisition, *Second Language Research, 23*(2): 215-42.

Tsimpli, I. M., & Mastropavlou, M. (2007). Feature interpretability in L2 acquisition and SLI: Greek clitics and determiners. In J. Liceras, H. Zobl, & H. Goodluck, (Eds.), *The Role of Formal Features in Second Language Acquisition* (pp. 143-183). New York: Routledge.

Unsworth, S. (2005). *Child L2, adult L2, child L1: Differences and similarities. A study on the acquisition of direct object scrambling in Dutch* (PhD dissertation). Utrecht University.

Unsworth, S. (2016). Early child L2 acquisition: Age or input effects? Neither, or both? *Journal of Child Language, 43*, 603-34.

Unsworth, S., Argyri, F., Cornips, L., Hulk, A., Sorace, A. & Tsimpli, I. M. (2014). The Role of Age of Onset and Input in Early Child Bilingualism in Greek and Dutch. *Applied Psycholinguistics, 35*(4): 765-805.

Unsworth, S., Hulk, A., & Marinis, T. (2011). Internal and external factors in child second language acquisition. (Introduction to spotlight issue of) *Linguistic Approaches to Bilingualism, 1,* 207-212.

Unsworth, S. & Hulk, A. (2009). Early successive bilingualism: Disentangling the relevant factors. Commentary on keynote by Jürgen Meisel. *Zeitschrift für Sprachwissenschaft, 28,* 69-77.

Vainikka, A. & Young-Scholten, M. (1994). Direct access to X'-Theory. Evidence from Korean and Turkish adult learners of German. In T. Hoekstra, B. Schwartz, (Eds.), *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins.

Vainikka, A. & Young-Scholten, M. (1996a). The early stages in adult L2 syntax: additional evidence from Romance speakers. *Second Language Research, 12,* 140-176.

Vainikka, A. & Young-Scholten, M. (1996b). Gradual development of L2 phrase structure. *Second Language Research, 12,* 7-39.

Vainikka, A. & Young-Scholten, M. (2007). Minimalism vs. organic grammar. In S. Karimi, V. Samiian, & W. Wilkins, (Eds.), *Phrasal and clausal architecture: syntactic derivation and interpretation. In honor of Joseph Emonds* (pp. 319-338). Amsterdam: John Benjamins.

Vainikka, A. & Young-Scholten, M. (2011). *The acquisition of German: introducing organic grammar.* Berlin: De Gruyter Mouton.

Vendler, Z. (1957). Verbs and times. *Philosophical Review, 56,* 143–60.

Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca Cornell University Press.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing, 17*, 65-83.

Wexler, K. (1990). *Optional infinitives, head movement and the economy of derivations in child grammar.* Paper presented at the Annual Meeting of the Society of Cognitive Science, MIT, Cambridge, Mass.

Wexler, K. (1992). *Optional infinitives, head movement and the economy of derivation in child grammar.* Occasional Papers #45. Center for Cognitive Science, MIT, Cambridge, Mass.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightfoot, & N. Hornstein, (Eds.), *Verb Movement* (pp. 305-350). Cambridge: Cambridge University Press.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua, 106*, 23-79.

Wexler, K. (2002). Lenneberg's dream: Learning, Normal Language Development and Specific Language Impairment. In J. Schaeffer and Y. Levy (eds.), *Towards a Definition of Specific Language Impairment* (pp. 11–62). Erlbaum.

White, L. (1989). *Universal grammar and second language acquisition*. Amsterdam: John Benjamins.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research, 7,* 133-161.

White, L. (1990/1991). The verb-movement parameter in second language acquisition. *Language Acquisition, 1,* 337–60.

White, L., Spada, N., Lightbown, P.M. & Ranta, L. (1991). Input enhancement and L2 question formation. *Applied Linguistics, 12*(4), 416-432.

White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

Whong-Barr, M., & Schwartz, B. D. (2002). Morphological and syntactic transfer in child L2 acquisition of the English dative alternation. *Studies in Second Language Acquisition, 24*(4), 579–616.

Yang, S., & Huang, Y. Y. (2004). The impact of the absence of grammatical tense in L1 on the acquisition of the tense-aspect system in L2. *International Review of Applied Linguistics in Language Teaching, 42*(1), 49-70.

Zdorenko, T. & Paradis, J. (2008). The acquisition of articles in child L2 English: Fluctuation, transfer or both? *Second Language Research, 24*, 227-250.

Zdorenko, T. & Paradis, J. (2011). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language*, *31*, 1–25.

Zhou, X. (2011). Early childhood education policy development in China. *International Journal of Child Care and Education Policy, 5*(1), 29-39.

Zobl, H. (1998). Representational changes: From listed representations to independent representations of verbal affixes. In M.-L. Beck (Ed.), *Morphology and its interfaces in second language knowledge* (pp. 339-372). Amsterdam: John Benjamins Publishing.

Zobl, H. & Liceras, J. (1994). Functional categories and acquisition orders. *Language Learning, 44*, 159–80.

APPENDIX A

A1. Ethics Approval letter

UNIVERSITY OF CAMBRIDGE
School of the Humanities and Social Sciences

Mrs Jill Noble
Ethics Committee Secretary

Athina Ntalli
Department of Theoretical and Applied Linguistics
Sidgwick Site
Cambridge

02 March 2017

Dear Athina

Ethical approval: Child Second Language Acquisition

The Chair of the Ethics Committee for the School of the Humanities and Social Sciences, acting on the Committee's behalf, has considered the documentation you provided, which followed the procedures concerning ethical approval of research.

I am able to inform you that approval, with respect to ethical considerations, has now been given to your project. Please note that this clearance is based on the documentation you have submitted. You must resubmit your application to the Ethics Committee should you subsequently make any substantive changes relating to matters reviewed by the Committee.

We are content for this letter to be forwarded to your grant sponsors, **EF- Education First.**

Yours sincerely

Jill Noble
Ethics Committee Secretary

cc Dr. Theodora Alexopoulou
Gabrielle Uncles
Elaine Schmidt

17 Mill Lane
Cambridge CB2 1RX

Tel: +44 (0) 1223 766238
Fax: +44 (0) 1223 760433
Email: cshssethics@admin.cam.ac.uk
www.cshss.cam.ac.uk

A.2.1 Consent form used prior to the amendment of the Data Protection Act.

UNIVERSITY OF CAMBRIDGE
Department of Theoretical and Applied Linguistics

Faculty of Modern and Medieval Languages
Sidgwick Avenue
Cambridge, CB3 9DA

an496@cam.ac.uk
+44 (0)7397312663

Study name or ID: Child Second Language Acquisition Research Project

Participant identification for this trial: .... (data to be retained anonymously)

CONSENT FORM

Name of Researcher: Athina Ntalli

Please initial box:

1. I confirm that I have read and understand the information letter dated 20/10/2017 for the above study and have had the opportunity to ask questions. I have understood them as given to me by the experimenter and/or read the paper version. ☐

2. I understand that my child's participation is voluntary and that he / she is free to withdraw at any time, without giving any reason, without his / her rights being affected. ☐

3. I understand that the data that will be generated in the study can be used in analyses, publications and teaching by researchers of the Child Second Language Acquisition project. I understand that audio recordings and their transcripts will be entered to the EF Cambridge Open Language Database (EFCAMDAT), a database for language research studies which is available to researchers worldwide. I understand that all data entered in EFCAMDAT do not contain the personal data of the EF students and that the data will be analysed and reported in an anonymous way. ☐

4. I agree that my child can take part in the above study. ☐

The Child Second Language Acquisition project has received ethical approval from the Humanities and Social Sciences Research Ethics Committee of the University of Cambridge.

Name of child’s parent	Date	Signature
Name of child participant	Date	Signature
Name of researcher	Date	Signature

A.2.2 Consent form used after the amendment of the Data Protection Act.

CONSENT FORM

Project title: Child Second Language Acquisition Research Project

Research team: Athina Ntalli, Dr Theodora Alexopoulou, Dr Henriette Hendriks

If you have any questions, please contact Athina Ntalli, at an496@cam.ac.uk

☐ I confirm that I have read and understand the information sheet dated 01/10/2018 for the above-mentioned study and have had the opportunity to ask questions.

☐ I understand that my child's participation is voluntary and that s/he is free to withdraw at any time, without giving any reason, and without their rights being affected.

☐ I understand that any data that are collected will be used and stored anonymously, in accordance with the Data Protection Act. Results are normally presented in terms of groups of individuals. If any individual data were presented, the data would be completely anonymous, without any means of identifying the individuals involved.

☐ I understand that these data may be used in analyses, publications, and conference presentations by researchers at the University of Cambridge and their collaborators at other research institutions. I understand that audio recordings and their transcripts will be entered to the EF Cambridge Open Language Database (EFCAMDAT), a database for language research studies which is available to researchers worldwide. I understand that all data entered in EFCAMDAT do not contain the personal data of the EF students. I give permission for these individuals to have access to these data.

☐ I understand that personal information (such as language background, age and gender) will be collected as part of this research. Full data will only be accessible to the research team. However, anonymised data may be used in analyses, publications and conference presentations. For full details on how we use your/your child's personal information, see https://www.information-compliance.admin.cam.ac.uk/data-protection/research-participant-data

☐ I have been given a copy of this form to keep.

I ______________________________________ agree my child to participate in the above-mentioned study run by Athina Ntalli, a PhD student at the Faculty of MML at the University of Cambridge.

____________________________________ ____________________

Date	Signature of participant	
Name of researcher	Date	Signature of researcher

A.3 Parent information sheet

Cambridge, 01-10-2018

INFORMATION LETTER
CHILD SECOND LANGUAGE ACQUISITION RESEARCH PROJECT

Dear parent,

I am a researcher at the University of Cambridge investigating how children learn English. We have a research collaboration with EF-Education First who have kindly opened their schools for us to collect data for our study. Below please find more information for our project. I hope that you will be able to support it and agree that your child participates in our study.

Who is participating?
Russian-speaking children who learn English at the EF (Education First) schools. Children should be either 9 or 12 years old and should have started learning English at 4 and 7 respectively.

Where will it take place?
The testing will take place at your child's EF school.

What will happen during the meeting session?
I will sit with the child in a quiet room and I will explain to the child what the tasks are about. I will tell a story to the child showing him/her pictures and I will ask the child to retell the story using their own words. I will also ask the child to repeat some sentences or to answer some questions about some pictures. The materials are children friendly and kids enjoy them. We will not evaluate the children in any way. The whole process needs to be audio recorded (no video recording) for later use and for research only purposes. It lasts approximately 45 minutes.

What will happen with the data?
Our aim is to understand the different ways in which children learn languages. Data collection is entirely anonymous and confidential. The data will become part of the EF Cambridge Open Language Database (EFCAMDAT), a database for language research studies TAL has been constructing in partnership with EF. The database is available to researchers worldwide investigating learning of English as a foreign language. All data entered in EFCAMDAT do not contain the personal data of the EF students. The data will be analysed and reported in a strictly anonymous way. Consent forms are only accessible to the researchers involved and will be used and stored in compliance with the Data Protection Act.

Participation in our research is entirely voluntary. Children can drop out from sessions if they wish to at any point.

Ethical review of the study

The project has received ethical approval from the Ethics Committee of the University of Cambridge.

Your child’s participation in this study is extremely valuable for us as this important research cannot take place without you. Understanding second language acquisition will further enable us to assist children’s learning and may have implications for educational policies and teaching methods.

Do contact me if you have questions about the process or if you would like to know more about the research. Thank you for your support to our project and look forward to meeting your son/daughter.

Yours sincerely,

Athina Ntalli
Email: an496@cam.ac.uk
Tel: (+44) 07397312663

A.4 Background questionnaire

UNIVERSITY OF CAMBRIDGE
Department of Theoretical and Applied Linguistics

Language Background Questionnaire

Dear Parent, we highly appreciate your contribution to this project and we value your answers. It is very important to us that you answer as objectively as possible because this will help us extract reliable results. If you are not sure what to answer to a certain question, please ask us.

1. General Information about the child

1.1. Name of child: ________________________________

1.2. Birth Date: ___________________________________

1.3. Gender: ______________________________________

2. Age of onset in learning English

2.1. At what age did your child start learning English at EF? ________________________________

2.2. At what age did your child start learning English at his/her day school? ______________________

2.3. Did your child attend another English language school before starting at EF? __________________

If yes, at what age did he/she start there and what kind of school was it (bilingual, private afternoon school, etc.)? ____

3. Exposure and use of English

3.1. How many years has your child been learning English in total? ____________________________

3.2. How many hours per week does your child attend English classes at his/her day school? ____________

3.3. How many hours per week does your child attend English classes at EF schools? _________________

3.4. How many hours per week does your child spend on English homework? _____________________

3.5. If your child uses electronic devices, how many hours per week does your child spend on them hearing or using the English language, e.g. playing games, exchanging messages (email, chat), reading websites, watching videos or listening to songs?____

3.6.1. Is there any other case, in which your child is exposed to English (e.g. radio, TV)? ______________

3.6.2. If yes, how many hours per week? ____

3.7. Have you been travelling to any English-speaking country? If yes, how often and how much time do you spend there? ____

4. Languages at home

4.1. Which is your child's native language(s)? ____

4.2. Which languages or dialects do you speak at home? ____

4.3. Which is the mother's native language(s)? ____

4.4. Which is the father's native language(s)? ____

4.5. Do you ever speak English to your child? If yes, how often? ______________________________________

5. Motivation / Attitudes

5.1. How important is it for you that your child learns English?

a. Very important b. Important c. Slightly important d. Not important

5.2. Does your child seem to enjoy learning English?

a. Very much b. Much c. Not much d. Not at all

6. Difficulties with your child's language development

6.1.1. Does your child have or has he/she ever had problems with the languages that he/she speaks? YES / NO

6.1.2. If yes, please describe:

____

____

7. Reading /Writing skills

7.1. At what age did your child learn to write and read in Chinese? ____________________________________

7.2. Does he/she know how to read and write in a language other than Chinese and English? ________________

8. Information about the family

8.1. Your level of education: ___________________________ (primary, secondary, upper secondary, professional training, undergraduate education, postgraduate education)

8.2. Your partner's level of education: ___________________________ (primary, secondary, upper secondary, professional training, undergraduate education, postgraduate education)

9. Additional information

9.1. Is there anything you would like to add or you consider we should take into account?

____

Thank you very much for your help and cooperation!

APPENDIX B

B.1 Recorded stories in retelling task

CAT (Total number of words: 178)

One day there was a playful cat who saw a yellow butterfly sitting on a bush. He leaped forward because he wanted to catch it. Meanwhile, a cheerful boy was coming back from fishing with a bucket and a ball in his hands. He looked at the cat chasing the butterfly.

The butterfly flew away quickly and the cat fell into the bush. He hurt himself and was very angry. The boy was so startled that the ball fell out of his hand. When he saw his ball rolling into the water, he cried: “Oh no, there goes my ball”. He was sad and wanted to get his ball back. Meanwhile, the cat noticed the boy’s bucket and thought: “I want to grab a fish.”
At the same time the boy began pulling his ball out of the water with his fishing rod. He did not notice that the cat had grabbed a fish. In the end, the cat was very pleased to eat such a tasty fish and the boy was happy to have his ball back.

Dog (Total number of words: 174)
One day there was a playful dog who saw a grey mouse sitting near a tree. He leaped forward because he wanted to catch it. Meanwhile, a cheerful boy was coming back from shopping with a bag and a balloon in his hands. He looked at the dog chasing the mouse.

The mouse ran away quickly and the dog bumped into the tree. He hurt himself and was very angry. The boy was so startled that the balloon slipped out of his hand. When he saw his balloon flying into the tree, he cried: “Oh no, there goes my balloon”. He was sad and wanted to get his balloon back. Meanwhile, the dog noticed the boy’s bag and thought: “I want to grab a sausage.”

At the same time the boy began pulling his balloon out of the tree. He did not notice that the dog had grabbed a sausage. In the end, the dog was very pleased to eat such a tasty sausage and the boy was happy to have his balloon back.

APPENDIX C

C.1 Data coding and scoring of the probes based on the TEGI manual

Table C.2: Coding/Scoring proposed by TEGI (3SG –s)

TEGI 3SG			
Structure attempted		**Structure not attempted**	
Correct	**Incorrect**	**Unscorable**	**No response**
3SG -*s* marking	Omission of –*s* marking	Other verb form/ tense (correct or not)	
*He play**s***	*He play*	*He is playing* *The dentist is make …* *The painter does drawing.* *The firefighter can put out the fire.*	*He …*

Table C.3: Coding/Scoring proposed by TEGI (Past)

TEGI PAST						
Structure attempted					**Structure not attempted**	
Regular verbs		**Irregular verbs**				
Correct	**Incorrect**	**Correct**		**Incorrect**	**Unscorable**	**No response**
Correctly inflected verb	Omission of –*ed*	Correct past form of irreg. verb	Overregu-larization	Incorrect form of irreg. verb	Any other form (e.g. another tense)	
*She play**ed***	*She play*	*She dug*	*She blowed (instead of blew)*	*She make*	*She cleans* *She was clean* **She making*	*She …*

C.2 Results of Chinese children on TEGI probes following the TEGI manual scoring method

Table C.4: Results on TEGI 3SG for Chinese learners following the TEGI manual guidelines

TEGI 3SG				
	STRUCTURE ATTEMPTED		**STRUCTURE NOT ATTEMPTED**	
	Correct 3SG (-*s*)	Incorrect 3SG (∅)	Unscorable	No response
9y.o. (n= 370)	38/187 20%	149/187 80%	180/370 **49%**	3/370 1%
12y.o. (n=34)	142/284 50%	142/284 50%	56/340 **16%**	0/340 0%

Table C.5: Results on TEGI Past for Chinese learners following the TEGI manual guidelines

TEGI PAST							
	Structure attempted					**Structure not attempted**	
	Regular verbs		**Irregular verbs**				
	Correct	Incorrect	Correct	Overregu-larization	Incorrect	Unscorable	No response
9y.o. (n=32)	90/143 63%	53/143 37%	41/129 32%	36/129 28%	52/129 40%	299/576 **52%**	5/576 1%
12y.o. (n= 34)	218/244 89.5%	26/244 10.5%	131/203 64.5%	52/203 25.5%	20/203 10%	165/612 **27%**	0/612 0%

APPENDIX D

D.1 Excluded verbs in narratives

Table 1: Excluded cases

Type	Example	Number of occurrences	Explanation
1. ‘wanna’ forms in 3SG contexts	*‘So, the cat saw the fish at the same time, so, he wanna eat it.’*	3	Forms like ‘wanna’ which do not show inflection while they may also be used in a formulaic fashion were left out.
2. Two verbs in a sentence	*‘There’s a boy comes …’*	12	In such cases, two interpretations can be made. It could be that children used the wrong form; in these cases, ‘comes’ instead of the target progressive participle. It could also be an omission of the complementiser ‘that’ which would make the sentence correct at least syntactically. I thus decided to include only the first verb in the scoring and exclude the second one from the analysis as it can be a wrong form instead of a progressive participle or a correct form if the complementiser is missing.
3. Agreement errors where the noun disagreed with the verb	*‘The big goat(’)s was looking for her.’*	26	In 26 examples, the child used a plural subject such as in ‘The big goats was looking for her.’ when the picture showed only

form / Ambiguous forms			one big goat. This might be a noun form error regarding number marking of the noun rather than a verb form error because the child could see the picture and could not have in mind a plural subject. Another possibility is to think that it could also be an auxiliary instead of a plural morpheme although there is another auxiliary following.
4. Agreement errors where the noun disagreed with the verb form.	*'The bird are happy.'*	2	Again, this can be considered as an error of number marking of the noun.
5. Use of a plural subject although there was only one entity in the picture.	*'mice was really quickly'*	4	I decided to exclude all these cases as some of them can be considered ambiguous in terms of type of error, and are in any case incorrect in terms of number marking on the noun.
6. Incorrect forms of main verbs	*'Their mother to catch food'* *'The mother to catching food'*	3	Incorrect use of 'to + infinitive /progressive participle' instead of a main verb. These instances would be categorised as 'Incorrect: other' but I decided to exclude them due to the very small number they appear.
7. Incorrect forms of main verbs	*'cames', 'bot', 'flow', 'saws'*	6	I decided to exclude these cases because they were ambiguous and incorrect.

8. Other syntactic issues	*'and the take it down'* *'and take out of the hands is balloon'*	2	This type of syntactic issues was excluded.
9. Verb + to + finite verb	*'she wants to he is on the grass and eating grass'*	1	Finite verbs after 'to' were excluded.
10. Incorrect strings of various tenses	*'boy is don't found his ball'* *'boy is don't found his ball under his hand'*	2	As these strings cannot be categorised for a particular tense marking and were just a couple of instances were left out.
11. First person singular subjects + main thematic verbs	*'I think', 'I don't know how to say that',*	18	First person singular subjects were a few while they may be formulaic expressions 'I think', 'I don't know how to say that', 'I can see', etc.
12. Third person plural subjects + main thematic verb	*'They walk …'*	46	The focus of the study is on 3SG contexts that can be assessed more reliably. Also, I cannot be sure about whether main thematic verbs following third plural subjects are inflected or bare as in present tense no inflection is used at all.
TOTAL		125	