



ACQUISITION OF THE VERBAL DOMAIN IN CHILD GREEK
EVIDENCE FROM A NEW CHILD GREEK CORPUS

Thesis submitted for the degree of Doctor of Philosophy
The School of Psychology and Clinical Language Sciences

Thomas Doukas

June 2011

ABSTRACT

The current thesis addresses the acquisition of the verbal domain in early Greek by exploring tense, finiteness, and subject-verb agreement based on samples from two monolingual children aged 1;7 – 2;11. The analyses of the data address two main theoretical accounts of language acquisition, namely, the generative approach and the usage-based approach. The results of the analyses, however, suggest that the latter approach did not provide sufficient empirical evidence to account for the data presented in this study.

Previous research suggested that sigmatic past in Greek is more prominent than non-sigmatic past, and therefore, its acquisition is subject to a dual mechanism. The results of the use of past tense suggest that sigmatic forms are used more often than non-sigmatic ones. A frequency analysis suggests that high frequency past tense forms in adults are used more often by children than low frequency ones.

Studies in child Greek proposed an early stage of development, during which children produce non-finite non-adult verbal forms, also referred to as the Root Infinitive stage. The data analysis show very few non-finite non-adult forms. These occur in children's speech only for a very short period at around the age of 2 years. The frequency analysis reveals that input does not relate to the production of RIs in children's speech.

Previous studies on the acquisition of verbal morphology showed that children's use of person and number markings are not productive and that children use mainly the 3rd singular. The subject-verb agreement analysis shows that error rates are low in children's speech and that subject-verb agreement is used productively from very early.

A frequency analysis shows that use of inflectional morphology is very similar between the two children but different to adults.

To conclude, this thesis provides new evidence for very early acquisition of the verbal domain in child Greek.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Thomas Doukas

ACKNOWLEDGEMENTS

Embarking on this journey there was only one aphorism in my mind: *reach what you cannot* (Nikos Kazantzakis — Report to Greco).

Completing a doctoral thesis is a herculean task. One's feelings are mixed; there is relief, happiness, sheer terror, anticipation, worry, and nostalgia. There's also gratitude and appreciation due, to many individuals, so this is an opportunity for me to thank all these people that have been, directly and indirectly, involved in this task.

First of all, I would like to express my gratitude to my supervisor Theo Marinis, for his support, patience and continuous advice and guidance for the past years. Theo's contribution to this task has been invaluable.

I am grateful to a number of linguists for discussing my ideas, and for entertaining my linguistic apprehensions in the course of this task, but mostly for their friendship; these are Evi Sifaki, Dina Haidou, Vina Tsakali, and Laura Mutti.

Special thanks are due to James Fenton for reading and commenting on some parts of this thesis, for coffees, cakes and great conversation by the river, and for being a true friend.

I would like to thank Maria's and Eve's families for allowing them to participate in this study. My time spent with Maria and Eve during the recordings was the most joyful part of this thesis.

Not all of my time was spent writing this thesis. Over the past 4 years I have had the pleasure of knowing and working with a number of people at Choice Support. I am grateful to Sarah Maguire, Pippa Gascoigne and Kim Arnold, for encouraging me, for being flexible with my research duties, and for listening to all my worries. Thanks for being there for me!

Special thanks are due to Andrew Moore for his psychological support, his belief in me, and his long-lasting patience and for being constantly beside me; *sine qua non!*

Last but not least, a huge thanks to my beloved family, my parents and my siblings, for their emotional and financial support, love and encouragement during the past years. Surely without their constant presence, belief in me and care I could not have succeeded on this project. I hope I am making you really proud!

This thesis is dedicated to Marina Tsakiridou, the converted *Hypatia*.

Table of Contents

| | |
|--|----|
| LIST OF TABLES | 10 |
| LIST OF FIGURES | 14 |
| ABBREVIATIONS | 15 |
| | |
| 1 INTRODUCTION | 17 |
| | |
| 2 THEORIES OF LANGUAGE ACQUISITION | 22 |
| 2.1 Intro..... | 22 |
| 2.2 How is language acquired? | 22 |
| 2.3 Generative approaches to language acquisition..... | 23 |
| 2.3.1 The logical problem of language acquisition..... | 23 |
| 2.3.2 Innateness..... | 24 |
| 2.3.3 The language faculty..... | 27 |
| 2.3.4 The role of input in UG..... | 28 |
| 2.3.5 The Principles and Parameters Theory | 31 |
| 2.3.6 Acquisition and Parameter Setting | 33 |
| 2.3.7 The minimalist perspective | 35 |
| 2.3.8 Maturation and Continuity accounts of Language Acquisition | 38 |
| 2.4 Usage-based accounts of Language Acquisition | 40 |
| 2.4.1 Role of input | 44 |
| 2.4.2 Verb Island Hypothesis..... | 46 |
| 2.4.3 Frequency..... | 48 |
| 2.4.4 Pre- proto- and morphology proper | 48 |
| 2.5 Summary | 50 |
| | |
| 3 METHODOLOGY | 52 |
| 3.1 Methods of Data Collection | 52 |
| 3.2 The data..... | 56 |
| 3.3 Maria's data | 57 |
| 3.3.1 The data collection procedure and transcription..... | 58 |
| 3.3.2 Coding..... | 59 |
| 3.4 Eve's data..... | 59 |
| 3.4.1 The data collection procedure..... | 61 |
| 3.4.2 Transcription | 61 |
| 3.4.3 Coding..... | 62 |
| 3.4.4 Other remarks on the coding..... | 73 |
| 3.4.5 Using the coding | 77 |
| 3.4.6 CLAN (Computerised Language Analyses) | 78 |
| 3.4.7 Frozen expressions..... | 80 |
| 3.4.8 Imitations and self-repetitions | 85 |
| 3.5 Stages of development: Age and MLU..... | 85 |
| 3.5.1 MLUw..... | 89 |
| 3.6 Contextual information | 91 |

| | | |
|-------|---|-----|
| 4 | PAST TENSE | 93 |
| 4.1 | Aims of the chapter | 93 |
| 4.2 | Past tense formation in Modern Greek | 95 |
| 4.3 | Aspect | 95 |
| 4.4 | Tense (past)..... | 97 |
| 4.5 | Linguistic accounts of past tense in Greek | 100 |
| 4.5.1 | Analysis 1 | 101 |
| 4.5.2 | Analysis 2 | 103 |
| 4.5.3 | Analysis 3 | 104 |
| 4.6 | Acquisition of past tense in Modern Greek | 110 |
| 4.6.1 | Stathopoulou (2009) | 111 |
| 4.6.2 | Delidaki and Varlokosta (2003)..... | 115 |
| 4.6.3 | Stavrakaki and Clahsen (2009)..... | 117 |
| 4.6.4 | Mastropavlou (2006)..... | 119 |
| 4.7 | Summary..... | 121 |
| 4.8 | Analyses..... | 123 |
| 4.8.1 | Sigmatic morphology..... | 123 |
| 4.8.2 | Productivity of verb inventory | 129 |
| 4.8.3 | MLUw..... | 131 |
| 4.8.4 | Augmentation and Regularity | 133 |
| 4.8.5 | Regularity..... | 138 |
| 4.9 | Trace-back procedure | 141 |
| 4.9.1 | Method..... | 141 |
| 4.9.2 | Frequency..... | 151 |
| 4.10 | Summary and discussion | 157 |
| 5 | ROOT INFINITIVES | 160 |
| 5.1 | Aims of the chapter | 160 |
| 5.2 | Introduction..... | 162 |
| 5.3 | Analyses on RIs with Continuity and Maturation accounts | 165 |
| 5.4 | Usage-based account of RIs..... | 169 |
| 5.4.1 | Simulation of RIs using MOSAIC..... | 170 |
| 5.4.2 | Kupisch and Rinke (2007) | 171 |
| 5.5 | Non-finite forms in <i>Modern Greek</i> | 172 |
| 5.5.1 | Verbal mood | 173 |
| 5.5.2 | Subjunctive and future | 173 |
| 5.5.3 | The Dependent..... | 175 |
| 5.5.4 | Imperatives..... | 176 |
| 5.5.5 | Person marking | 177 |
| 5.6 | Linguistic accounts of RIs in Greek | 178 |
| 5.6.1 | Modality in Modern Greek (Stephany 1981, 1986, 1995)..... | 178 |
| 5.6.2 | First approaches in the analysis of RIs – the -i form | 180 |
| 5.6.3 | Bare Perfectives (Hyams 2002, 2005, 2007) | 185 |
| 5.6.4 | New evidence for the non-finite forms (Varlokosta 2002, 2005)..... | 191 |
| 5.6.5 | Imperatives – RI analogue in Null Subject Languages..... | 194 |
| 5.6.6 | Semantic interpretation and modal properties of RIs and BPs | 196 |
| 5.6.7 | Person and Infinitive features and RI stage length continuum (Perales, Licerias and Bel)..... | 198 |

| | | |
|-------|---|-----|
| 5.7 | Summary | 201 |
| 5.8 | Analyses..... | 202 |
| 5.8.1 | Errors | 203 |
| 5.8.2 | Modal particles (na/tha) analysis | 212 |
| 5.8.3 | RIs | 220 |
| 5.8.4 | Other evidence for RIs | 234 |
| 5.8.5 | Verb typology and temporal interpretation of BPs..... | 245 |
| 5.8.6 | Modal interpretation of BPs..... | 247 |
| 5.8.7 | A note on usage-based analyses..... | 255 |
| 5.9 | Summary and discussion | 258 |
| 6 | PERSON AND NUMBER MARKING WITHIN THE VERBAL DOMAIN | 263 |
| 6.1 | Aims of the chapter | 263 |
| 6.2 | Introduction..... | 265 |
| 6.3 | Literature Review | 268 |
| 6.3.1 | Very Early Parameter Setting and Very Early Knowledge of Inflection | 268 |
| 6.3.2 | Usage-based models | 269 |
| 6.4 | Person and number in Modern Greek | 274 |
| 6.5 | Acquisition of verbal inflection in MG..... | 275 |
| 6.5.1 | Number and Person Markers (Katis) | 275 |
| 6.5.2 | Tsimpli | 277 |
| 6.5.3 | Christofidou and Stephany (2003)..... | 278 |
| 6.6 | Analyses..... | 280 |
| 6.6.1 | Person morphology | 280 |
| 6.6.2 | Error analysis | 293 |
| 6.6.3 | Error analysis by context | 295 |
| 6.7 | Productivity..... | 304 |
| 6.7.1 | Productivity of the use of person and number marking..... | 304 |
| 6.7.2 | Productivity in Maria's and Eve's speech | 306 |
| 6.7.3 | Frequency analysis..... | 313 |
| 6.8 | Summary and discussion | 315 |
| 7 | GENERAL DISCUSSION AND CONCLUSIONS..... | 319 |
| 7.1 | Past Tense | 320 |
| 7.2 | Root Infinitives | 326 |
| 7.3 | Number and Person marking | 332 |
| 7.4 | Limitations of the study and future research | 337 |
| 7.5 | Conclusion | 339 |
| | APPENDICES | 342 |
| | Appendix A: Are verbs islands? | 342 |
| | REFERENCES..... | 350 |

LIST OF TABLES

| | |
|---|-----|
| Table 2.1: Bundles of analyses related to the features of the verbal domain in MG | 37 |
| Table 3.1: Summary of children's general information..... | 52 |
| Table 3.2: Transliteration system used for the transcription of the data..... | 56 |
| Table 3.3: Maria's tape recordings' details | 58 |
| Table 3.4: Eve's tape recordings' details..... | 60 |
| Table 3.5: Examples of verb coding | 63 |
| Table 3.6: Conditions for formula identification | 81 |
| Table 4.1: Perfective-imperfective aspect distinction for verb <i>graf-o</i> (= write)..... | 96 |
| Table 4.2: Past (simple) tense formation and irregularity | 105 |
| Table 4.3: Summary of the accounts on MG past tense formation (features) | 108 |
| Table 4.4: Maria's sigmatic/non-sigmatic past tense forms | 126 |
| Table 4.5: Eve's sigmatic/non-sigmatic past tense forms..... | 128 |
| Table 4.6: Maria's sigmatic/non-sigmatic distribution by MLUw | 131 |
| Table 4.7: Eve's sigmatic/non-sigmatic distribution by MLUw | 133 |
| Table 4.8: Maria's [+/- augment] in past tense..... | 135 |
| Table 4.9: Eve's [+/- augment] in the past tense | 136 |
| Table 4.10: Maria's [+/-regular] in sigmatic and non-sigmatic past tense | 139 |
| Table 4.11: Eve's [+/-regular] in sigmatic and non-sigmatic past tense | 139 |
| Table 4.12: Ratio of analysed utterance between test recording and main corpus | 143 |
| Table 4.13: Maria's matched past tense in the main corpus and in adults' speech | 145 |
| Table 4.14: Eve's matched past tense in the main corpus and in the adults' speech.... | 146 |

| | |
|--|-----|
| Table 4.15: High frequency used past tense in Maria’s and Eve’s speech (divided into sigmatic and non-sigmatic)..... | 152 |
| Table 4.16: Low frequency use of past tense in Maria’s and Eve’s speech (divided in sigmatic and non-sigmatic)..... | 153 |
| Table 4.17: Frequencies of all past tense verbs from the Hellenic National Corpus in descending order (sigmatic in italics)..... | 154 |
| Table 5.1: Realis and irrealis opposition expressed in [+/-P] and [+/-R] features in different languages..... | 200 |
| Table 5.2: Maria’s overall error rate (verbal domain)..... | 203 |
| Table 5.3: Eve’s overall error rate (verbal domain)..... | 204 |
| Table 5.4: Janna’s correct and incorrect use of agreement..... | 205 |
| Table 5.5: Rate of agreement error in children’s languages..... | 206 |
| Table 5.6: Maria’s errors (by category)..... | 210 |
| Table 5.7: Eve’s errors (by category)..... | 211 |
| Table 5.8: Maria’s <i>na</i> -constructions with modal particle used and omitted..... | 213 |
| Table 5.9: Eve’s <i>na</i> -constructions with modal particle used and omitted..... | 214 |
| Table 5.10: Maria’s use of <i>na</i> -clauses..... | 216 |
| Table 5.11: Eve’s use of <i>na</i> -clauses..... | 216 |
| Table 5.12: Maria’s future constructions with particle <i>tha</i> used and omitted..... | 217 |
| Table 5.13: Eve’s future constructions with particle <i>tha</i> used and omitted..... | 218 |
| Table 5.14: Maria’s 3 rd singular with <i>-i</i> and other endings..... | 224 |
| Table 5.15: Eve’s 3 rd singular with <i>-i</i> and other endings..... | 226 |
| Table 5.16: Maria’s <i>-i</i> forms with aspectual stems..... | 228 |
| Table 5.17: Eve’s <i>-i</i> forms with aspectual stems..... | 229 |

| | |
|---|-----|
| Table 5.18: Distribution of BPs in Maria's data | 231 |
| Table 5.19: Distribution of BPs in Eve's data | 232 |
| Table 5.20: Stem distribution in [+past] verbs in Maria's and Eve's corpora..... | 237 |
| Table 5.21: Clitic object omission in 3sg vs. other persons in Maria's data | 242 |
| Table 5.22: Clitic object omission in 3sg vs. other persons in Eve's data..... | 242 |
| Table 5.23: Clitic objects in 3sg <i>-i</i> forms (BPs and non) in Maria's data | 243 |
| Table 5.24: Clitic objects in 3sg <i>-i</i> forms (BPs and non) in Eve's data..... | 243 |
| Table 5.25: Temporal reference of Maria's and Eve's BPs..... | 246 |
| Table 6.1: Rate of Agreement error in children's languages | 267 |
| Table 6.2: Maria's accuracy of person morphology of the verbal paradigm in the singular (errors of form) | 281 |
| Table 6.3: Maria's accuracy of person morphology of the verbal paradigm in the plural (errors of form) | 282 |
| Table 6.4: Eve's accuracy of person morphology of the verbal paradigm in the singular (errors of form) | 287 |
| Table 6.5: Eve's accuracy of person morphology of the verbal paradigm in the plural (errors of form) | 288 |
| Table 6.6: Maria's Agreement Errors | 293 |
| Table 6.7: Eve's Agreement Errors | 294 |
| Table 6.8: Maria's distribution of Agreement errors (context errors) | 295 |
| Table 6.9: Eve's distribution of Agreement errors (context errors) | 296 |
| Table 6.10: Maria's and adult use of singular inflectional endings..... | 314 |
| Table 6.11: Eve's and adult use of singular inflectional endings | 314 |

| | |
|--|-----|
| Table 7.1: Analyses related to the features of the verbal domain in MG examined in this thesis | 319 |
|--|-----|

LIST OF FIGURES

| | |
|--|-----|
| Figure 3.1: Maria's MLUw..... | 90 |
| Figure 3.2: Eve's MLUw | 90 |
| Figure 5.1: The RI stage length <i>continuum</i> | 201 |

ABBREVIATIONS

| | | |
|------------------|---|--|
| ACC | : | Accusative |
| AdjP | : | Adjective Phrase |
| AGR | : | Agreement |
| AGRP | : | Agreement Phrase |
| BLF | : | Broad Language Faculty |
| BPs | : | Bare Perfectives |
| CHAT | : | Codes of the Human Analysis of Transcripts |
| CHILDES | : | Child Language Data Exchange System |
| CLAN | : | Computerised Language Analyses |
| CP | : | Complementiser Phrase |
| DAR | : | Default Anchoring Requirement |
| DEP | : | Dependent |
| DIM | : | Diminutive |
| DP | : | Determiner phrase |
| DS | : | Down's Syndrome |
| <i>e</i> | : | Event |
| EMC | : | Early Morphosyntactic Convergence |
| EC | : | Eventivity Constraint |
| FREQ | : | Frequency |
| FCH | : | Full Competence Hypothesis |
| FM | : | Functional Module |
| FUT | : | Future |
| IAH | : | Imperative Analogue Hypothesis |
| IQ | : | Intelligence quotient |
| IMP | : | Imperative |
| IMPERF | : | Imperfective (aspect) |
| INFL | : | Inflection |
| IPA | : | International Phonetic Alphabet |
| IP | : | Inflectional Phrase |
| LAD | : | Language Acquisition Device |
| LF | : | Logical Form |
| MG | : | Modern Greek |
| MLU | : | Mean Length of Utterance |
| MLU _w | : | Mean Length of Utterance in words |
| MLT | : | Mean Length of Utterance in a turn |
| MOSAIC | : | Model of Syntax Acquisition In Children |
| MRE | : | Modal Reference Effect |
| MSP | : | Mean Size of Paradigm |
| NLF | : | Narrow Language Faculty |
| NOM | : | Nominative |
| NP | : | Noun Phrase |
| OI | : | Optional Infinitives |
| PBF | : | Percentage of Base Form |
| PERF | : | Perfective (aspect) |
| PF | : | Phonetic Form |
| PFI | : | Principle of Full Interpretation |

| | | |
|----------------|---|--|
| P&P | : | Principles and Parameters |
| PP | : | Prepositional Phrase |
| [+/-P] | : | Person (feature) |
| RI | : | Root Infinitives |
| [+/-R] | : | Infinitive (feature) |
| SBH | : | Structure Building Hypothesis |
| SLI | : | Specific Language Impairment |
| SUBJ | : | Subjunctive |
| SVO | : | Subject Verb Object |
| S ₀ | : | Initial state |
| S _s | : | Steady state |
| [+/-SS] | : | Stress-shift feature |
| TAR | : | Temporal Anchoring Requirement |
| TP | : | Tense Phrase |
| UG | : | Universal Grammar |
| UT | : | Utterance Time |
| VEPS | : | Very Early Parameter Setting |
| VEKI | : | Very Early Knowledge of Inflection |
| VP | : | Verb Phrase |
| VVR | : | Varlokosta, Vainikka and Rohrbacher (1996, 1998) |
| WS | : | Williams Syndrome |

1 INTRODUCTION

Research in language acquisition in many languages has demonstrated that children aged 2-3 years have already acquired different aspects of grammar. Nonetheless, it has been claimed that children's language and grammar is not identical to the one used by adults. Therefore, children's languages, is claimed to be subject to restructuring in order to comply with the rules of the adult model.

The chief quest of language acquisition research is explaining the difference between infant grammar on one hand, and that of adults on the other, and the process of development from this early language to the adult target. Such research questions have given ground for the development of two different theories, i.e. the generative approach and the constructivist or usage-based approach. The generative approach to language acquisition maintains that language is a cognitive faculty where the grammar can be described as an internalized set of formal rules and principles which generate sentences in the language. According to the usage-based theories, language is considered to be acquired from the input that children are exposed to and patterns are inferred from the data.

The main focus of this thesis is to investigate the acquisition and development of some components of the verbal domain in child Greek through the analysis of spontaneous speech data collection. The analyses of the verbal components will examine features of both theories. The results of the analyses will be discussed within the frameworks provided by the theories mentioned above.

The verbal components under investigation are tense, inflection and subject-verb agreement.

The study of past tense explores the status of acquisition of past tense in Modern Greek addressing two questions: firstly, from a generative point of view, the difference between sigmatic and non-sigmatic past tense in children's early speech (regular vs. irregular formation) and secondly, from a usage-based point of view, the children's production of past tense in relation to their own previous production (trace back) and in relation to the adults' production. The predictions for the first question, related to the morphological features of past tense, are that sigmatic past is acquired and used more in children's speech than non-sigmatic past. The predictions for the second question, related to frequency, are that children's frequently used past tense can be traced back in their own previous production and also that these items are frequently used in the adult language.

The study of inflection is related to the use of finite and non-finite verbal forms in child Greek. Based on previous research on the presence of non-finite non-adult verbal forms in children's speech (Root Infinitives) the analysis addresses two questions: from a generative point of view, it looks at the use of non-finite verbal forms in children's production and the existence of two different stages in the acquisition of the verbal system in Modern Greek; from a usage-based viewpoint, related to frequency, it is looking at children's non-finite non-adult constructions as being the result of frequency of some adult structures or reduced productions of some adult utterances. Finally, the status of morphological ambiguity in Modern Greek is examined, being a proposed cause of children's non-adult production. The predictions for the first question, based on a series of analyses reproduced from previous research, are that children's language develops in two stages; during the first stage, non-finite non-adult forms are overused while during stage two these forms are reduced and soon disappear. The

predictions to the second set of questions are that frequently used items in the adults' language will be frequently used in children's speech too and that the presence of non-finite non-adult forms is the result of some language specific morphological ambiguity.

The study of subject-verb agreement focuses on the development of person and number inflectional features in Maria's and Eve's speech production addressing the following questions: from a generative point of view, it looks at the status of person and number morphology in terms of emergence, distinction, and substitution during the process of acquisition and secondly at the productive use of these morphological features; from a usage-based viewpoint, it looks at the relation between children's frequently used items and the adult pattern of frequency. The predictions for the first set of questions are that 3rd person is acquired before 1st and 2nd, and singular before plural; also that children substitute 1st and 2nd person with 3rd and that the use of person and number morphology is not productive from the onset. The predictions for the second set of questions are that children use more frequently verb forms used frequently in the adult language they are exposed to.

This study was motivated by the lack of research that provides analyses for both theoretical frameworks, i.e. generative and usage-based. Research in language acquisition and more specifically in Modern Greek focuses on one or the other theory, thus automatically forsaking the other theory. There are not studies of child Greek language acquisition that attempt to analyse data based on elements from both theories. The analyses further investigate questions about the acquisition of tense, inflection and subject-verb agreement raised in previous research in Modern Greek, as discussed in each chapter.

A second motivation of this thesis was the limited data in child Greek available for examining the verbal domain from a developmental point of view. As discussed in each of the chapters on the analyses (chapters, 4, 5, and 6), to date there is only one set of longitudinal data in Modern Greek publicly available (cf. Stephany Corpus in CHILDES) dated from the early 70s. Two further collections of spontaneous data are not available online (cf. Christofidou corpus and Katis coprus).

A series of experimental data collected via elicitation tasks has also been used in research for child Greek (Mastropavlou, 2006; Stathopoulou, 2009; Stavrakaki and Clahsen, 2009). These studies often, if not always, used very different methodologies and they focused in very different grammatical phenomena. As a result, the available results and findings cannot be used for comparative analyses. The data also cannot provide any insight from a developmental point of view and the children tested in these studies are often older in age.

The collection of the data used in this thesis covers an age range from 1;7 to 2;11 of two Greek speaking children, namely, Maria and Eve. Therefore the data discussed are produced by children of younger age than the ones analysed in previous studies, also providing a greater age range that allows developmental analyses.

To summarise, this thesis contributes to the field of language acquisition firstly, by providing a comprehensive analyses of the two main theoretical frameworks i.e. generative and usage-based and secondly by using new data of spontaneous speech that cover a greater age range than previously studied and allow for a developmental analysis of early Greek.

The structure of the thesis is as follows:

Chapter 2 introduces the theories of language acquisition used in this study. This chapter also aims to provide a brief historical outline of theories that constituted the basis and foundation of modern linguistic theory, albeit not used directly in this study.

Chapter 3 outlines the methodology of this study. It describes how the data were collected, the transcription, coding scheme and the way the analyses of the data were conducted. Methodological problems encountered during the transcription and analyses of the data are also discussed.

Chapter 4 investigates the status of acquisition of the past tense in child Greek based on the analysis of spontaneous speech production of Maria and Eve. This chapter focuses on tense and aspect and, in particular past vs. non-past tense and perfective vs. imperfective aspect distinctions.

Chapter 5 examines the status of non-finite forms in early Greek. The aim of this chapter is to discuss previous research on RI stage in children's speech and subsequently analyse Maria's and Eve's data against the proposed evidence.

Chapter 6 studies the development of person and number verb morphology and subject-verb agreement. The aim of this chapter is to consider previous theoretical and empirical data about the acquisition of agreement, and thereafter, to examine the acquisition of person and number morphology in Maria's and Eve's data within the proposed framework.

Finally, chapter 7 provides a general discussion by summarising the results of the analyses carried out in chapters 4, 5, and 6 and by providing an overview of the main points of discussion from each chapter. This chapter also points out the limitations of this study and the issues that are left open for further research.

2 THEORIES OF LANGUAGE ACQUISITION

2.1 Intro

The study of language acquisition raises three fundamental questions:

- a) What does language knowledge consist of?
- b) How do children acquire language?
- c) How do children use this language and how does this language grow in given stages of the acquisition process?

In recent years, a variety of theories of language acquisition were developed in order to answer these questions.

The aim of this chapter is to explore those theories of language acquisition used in the consecutive chapters of this thesis. This chapter also aims to provide a brief historical outline of theories that constituted the basis and foundation of modern linguistic theory, albeit not used directly in this study.

2.2 How is language acquired?

There are two general approaches to the acquisition of language, i.e. the constructivist or usage-based approach and the generative approach.

According to the usage-based approach, language is considered to be acquired from the input that children are exposed to and patterns are inferred from the data. Generative approaches to language acquisition suggest that language is a cognitive faculty where the grammar can be described as an internalised set of formal rules and principles which generate sentences in the language.

2.3 Generative approaches to language acquisition

Generative approaches to language acquisition propose that there is a strong innate component used for language acquisition. Generativists make two main assumptions for language acquisition: firstly, the acquisition of language structure is explained through mechanisms inherent to the child (Chomsky, 1986; Goodluck, 1991; Atkinson, 1992). Chomsky (1986) proposed that human beings possess a unique language faculty that is responsible for the use and acquisition of language. What follows from this assumption is that parental input is important, but minimal, e.g. parents do not provide information about ungrammaticality of an utterance, and thus, children do not have access to negative evidence (Chomsky, 1984; Crain and Thornton, 1998; Valian, 2009). And secondly, children's language may appear to be different from the adult one, but with respect to the structure of the language system, this is said to be basically the same (Chomsky, 1988; Goodluck, 1991; Atkinson, 1992).

2.3.1 The logical problem of language acquisition

The problem in explaining how children acquire language is related to linguistic creativity or the potential infinite use of linguistic constructions. This problem, also known as Plato's problem or the poverty of the stimulus, claims that the nature of language knowledge is such that it could not be acquired from the actual samples of language available to the human child.

The poverty of the stimulus argument is founded on the following premises:

- Languages are complicated systems and thus necessitate complex descriptive devices.
- A native speaker of a particular language knows a particular aspect of this system e.g. syntax
 - Syntax could not have been acquired from the language input typically available to children
 - Syntax is built-in to the mind

If these premises are correct, then the question that arises is how children learn these aspects of language apparently effortlessly at an early age? In essence, the argument from the poverty of stimulus starts with the idea that all native speakers of a given language know some kind of abstract property of this language which cannot be induced from the evidence available from the input, the so called positive evidence (Guasti, 2002)

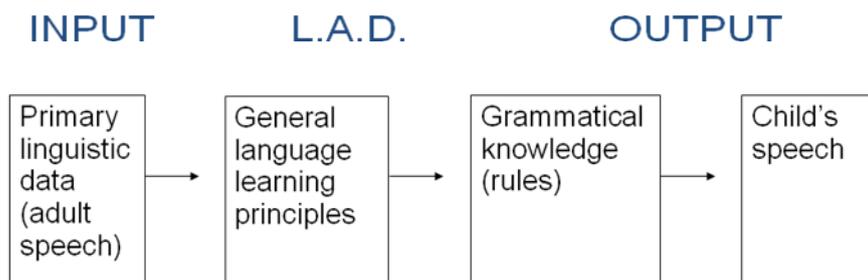
The solution, proposed by Chomsky, is to remove the complexities from individual grammars and *attribute them to the genetic endowment of the child* (Chomsky, 1999)

2.3.2 Innateness

In the 1960s, arising out of the generative framework of language acquisition and due to the limitations of previous theories, such as imitation and reinforcement, an alternative proposal argued that children must be born with an innate capacity for language development: the human brain is ‘ready’ for language, in the sense that when children are exposed to language, certain general principles for discovering or structuring

language automatically begin to operate (Chomsky, 1965). These principles constitute the child's 'language acquisition device' (LAD).

According to Chomsky (*ibid.*) children use their LAD to make sense of the utterances heard around them, and thus, assumptions about the grammar of the language they are exposed to – e.g. how sentences are constructed – derive from the 'primary linguistic data'. This knowledge is then used to produce sentences that, after a process of trial and error, comply with the generalisations and rules governing the way sentences are formed in the adult language. This sequence of events can be summarised in the following diagram:



The notion of LAD, originally formulated by Chomsky (1984, 1999), is necessary for explaining the remarkable speed with which children learn to speak, and the considerable similarity in the way grammatical patterns are acquired across different children and languages from the input.

Adult speech, it is argued, is too complex and unsystematic to provide alone the means for children to work out the regularities of language by themselves. It has proved difficult to formulate the detailed properties of LAD in an uncontroversial manner, in

the light of the changes in generative linguistic theory that have taken place in recent years.

Chomsky (1981, 1986) consequently, posited the theory of Universal Grammar (hereafter UG) as the set of constraints which characterises humans' genetically determined language faculty. UG is different from LAD in that it no longer suggests access to predefined grammars and it also reevaluates the role of primary input in acquisition of language (Lust, 2006).

Universal Grammar is the theory of possible human internal-languages i.e. a theory of the human biological endowment for language prior to experience (Chomsky, 1976, 1981). According to UG all humans are innately endowed with a linguistic system and therefore share the same linguistic knowledge regardless of which language they learn (Chomsky, 1976, 1981).

Innately biological properties are specified empirically when manifesting certain relevant characteristics. Similarly, innate linguistic property could not be innately specified without empirical evidence.

Evidence for innateness is provided by three main characteristics:

- a) Absence of input evidence, i.e. when a property appears in the absence of decisive evidence from the environment, i.e. children adhere to a linguistic principle for which there is no corresponding evidence in the environment, then the principle is likely to be innately specified (Valian, 1999; 2009). Children exposed to different input, although of the same language, adhere to a particular linguistic principle; such principle is therefore innately specified (Lust, 2006).
- b) Universality of linguistic principle: if a linguistic principle is part of the human biological blueprint for language growth, then it is expected to be manifested by

children learning any natural language – hence the name Universal Grammar (Guasti, 2002; Valian, 2009).

- c) Early emergence of linguistic principles is also suggested to provide evidence for innateness (Guasti, 2002; Lust, 2006). The early emergence of complex linguistic principles provides counter-evidence for learning-theoretic accounts of language acquisition.

These three requirements – that innate principles a) appear without decisive evidence from the environment, b) are universal, and c) emerge early – will be referred to as the *Innateness Hypothesis*: ‘it must be that the mind/brain provides a way to identify and extract the relevant information by means of mechanisms of some sort that are part of its biologically determined resources’ (Chomsky, 1988: 15).

2.3.3 *The language faculty*

Language faculty or instinct or capacity is the psychological ability of humans to acquire, produce and understand a particular language (and language in general) (Carnie 2002: 20). Our competence or knowledge of language also referred to as *internal language* is represented in the mind as the *grammar* of the particular language the speaker has acquired (Chomsky, 1986). The grammar in this sense includes all aspects of the linguistic structure of the language: phonetics, phonology, morphology, lexicon, syntax, semantics, and certain aspects of pragmatics.

The way this knowledge of language is put to use is referred to as the *external language* or performance and represents the processing mechanisms, e.g. language in context, speech errors etc. (Chomsky, 1986).

The nature of the language faculty is therefore twofold, namely, the Narrow Language Faculty (NLF) that includes only recursion and the Broad Language Faculty (BLF) which includes sensor motor, conceptual and grammatical components; most components are shared to different degrees with other species (Hauser et al., 2002).

2.3.4 *The role of input in UG*

As a consequence of the innateness hypothesis, for many years, the importance of the language used by adults (especially mothers) to children was minimised. Studies of ‘motherese’, as it came to be called in the 1970s¹, showed that maternal input is by no means as complex and fragmentary as proponents of innateness theory claimed it to be (Snow, 1994). Many parents do not talk to their children in the same way they talk to other adults (child oriented speech vs. adult oriented speech); they rather appear to be adapting their language to give the child maximum opportunity to interact and learn.

Several of these adaptations have been noted; the utterances are considerably shorter and simplified, especially with respect to their grammar and meaning (Harley, 2001). There is a more restricted range of sentence patterns, and a frequent use of sentence ‘frames’, such as *Where’s-?* or *That’s a-*. The meanings are predominantly ‘concrete’, relating to the situation in which mother and child are acting (Ferguson, 1978; Dockrell et al., 1999). There are several features whose purpose seems to be clarification. Extra information is provided that would be considered unnecessary when talking to other adults. Sentences are expanded and paraphrased and may be repeated several times. The speed of speaking is much slower than that used to other adults.

¹ Now called *child-directed speech*

There is an expressive, or affective, element in motherese, shown by the use of special words or sounds (Harley, *ibid.*). The use of diminutive or reduplicative words (e.g. *doggie*, *choo-choo*) is a commonplace. English makes particular use of a *y/ie* ending, and similar forms have been noted in several other language, such as Japanese – *ko*, Gilyak² –*k/-q*, Greek –*aki*. Occasionally, totally different words will be used, e.g. *bunny* for ‘rabbit’.

The prosodic modifications in parents’ speech to infants was investigated in the cross-linguistic study by Fernald et al. (1989) for French, Italian, German, Japanese, British English, and American English. The results showed that all parents across languages used common patterns of intonation, shorter utterances, and longer pauses in infant-directed speech than in adult-directed speech (Fernald, et al., *ibid.*). Some of the features also seem to function as ways of holding the child’s attention, or of identifying particular words and sounds and to facilitate speech perception and language comprehension (Fernald et al., *ibid.*).

Mothers also devote a great deal of time to obtaining feedback from their children, especially in the first three years (Newport et al., 1977; Ferguson, 1978). Their speech contains a very high frequency of question forms, and many utterances have a high rising intonation (*yes?*, *all right?*). These adaptations are important ways of establishing and maintaining meaningful communication with the child, as they can be found in the earliest mother-child interactions.

Although typically slower, clearer and with highlighted intonation, Fromkin et al., (2003) maintain that child directed speech is not necessarily syntactically simple; it

² A language of Russian Federation (Asia)

contains a variety of complex constructions such as questions; embedded clauses, imperatives, negation etc.

It has even been suggested that these features are universal, but this claim is premature in the absence of empirical studies, and there is already some counter-evidence from other cultures – several of these features are lacking in the Samoan and Quiche Mayan community, for instance (Pye, 1986). However, the highly structured character of maternal input is not in doubt, and its possible influence on the course of language acquisition is taken very seriously in recent research.

There is controversy, however, on the effect that child directed speech has on language development and the correlations between the features of motherese and the subsequent emergence of these features in child speech. Newport et al. (1977) tested the motherese and the Fine Tuning hypothesis³ (cf. Valian, 1999) on fifteen American mothers and their results showed that there is no direct correlation between the children's speech and the properties of the maternal speech.

Other studies have found occasional correlations⁴ between specific structures, e.g. Howe (1980) showed that children appear to have a more rapid linguistic development when their mothers asked more questions and gave them more extensive replies to their questions.

The use of different research methodologies complicates matters further, but what is now certain is that the nature and frequency of linguistic features in maternal input can no longer be neglected in devising theories of language acquisition.

³ When parents are acting as teachers, then as their speech grows in complexity, so should the child's, i.e. the input structures are very closely tailored to the needs of the child.

⁴ The correlation between input and production from a usage-based point of view is discussed in section 2.4.1.

2.3.5 *The Principles and Parameters Theory*

The Principle and Parameter theory (henceforth P&P), devised by Chomsky over the past two decades, maintains that universal properties of natural language grammars reflect the operation of a set of universal grammatical principles (Chomsky, 1981; Chomsky and Lasnik, 1993). According to the P&P model there is a set of universal principles which account for language structure in general, and a set of finite principles which capture cross-linguistic variation; grammatical principles and therefore grammatical differences between languages can be accounted for in terms of a restricted set of parameters (Chomsky, *ibid.*).

Principles describe potentially universal properties of grammatical operations or structures. The terms *conditions* and *constraints* are also used with much the same meaning. Potential principles of Universal Grammar include the structure dependence principle, the head movement constraint, the shortest movement principle, the economy principle, etc.

Parameters account for the grammatical variation between different languages or different varieties of the same language e.g. the null subject parameter, head parameter, wh-parameter.

According to the Principles and Parameters Model, UG comprises of two elements (Chomsky, 1981):

- a. A set of absolute universals, notions and principles which do not vary from one language to the next and,
- b. Language-specific properties which are not fully determined by UG but which vary cross-linguistically. For these properties a range of choices is made available by UG.

2.3.5.1 *Old and new theory*

At first, parameters were associated directly with the principles of UG (or vice versa) e.g. the Extended Projection Principle which guarantees that all clauses have subjects, had an associated parameter – the pro-drop or null-subject parameter – determining whether the subject could be empty (as in Greek and Italian), or was necessarily overt (as in English and French).

Parameters were set for the language as a whole. The major advantage of this was the so called “cascade effects” (when the child learns A automatically knows B and C). There were, however, major dissociations and inconsistencies where it looked as if a parameter had to be set in two ways in the same language e.g. German and Amharic⁵ have AdjP with both head-complement and complement-head sequences (a language could be only ‘head first’ or ‘head final’ in all its structural phrases).

To account for the contradictions of the original theory (Chomsky, 1981), newer versions were proposed (Borer 1984, Chomsky 1995) where parameterisation can be restricted to the lexicon and subsequently, parametric variation has been constrained to refer only to functional categories, and within that only to the value of a specific feature [+/- strong]. Chomsky (1995) suggested that if parameters fractionate and the cascade effects disappear, then the domain of any particular parameter is diminished, and with it construction-particular rules are eliminated.

⁵ A language of Ethiopia

2.3.6 *Acquisition and Parameter Setting*

As noted previously, the language faculty contains linguistic components that allow children to acquire and develop language effortlessly. The process of acquisition comprises of various stages starting with the initial state S_0 , i.e. the onset of language acquisition and resulting to the steady state S_S , i.e. the target grammar. Children go through a variety of stages between the initial and the steady state, i.e. S_1 , S_2 , S_3 , and so on.

The parameters can be thought of as built-in switches, each to be turned to suit the language that the child is exposed to. The transition from the initial state S_0 to the steady state S_S is a matter of setting the switches. The setting of switches is triggered by evidence from the input. Children must be learning from positive evidence alone or from indirect negative evidence. Hearing a few sentences is sufficient to set the parameter one way or another.

Three logical possibilities can be distinguished for parameters in the initial state S_0 :

- i. The switch is in a neutral position; children are equally prepared for setting the parameter to either [+] or [-], e.g. in the case of the pro-drop parameter children could choose either pro-drop or non pro-drop. In this case the interim stages in children's development of grammar might have either setting for pro-drop.
- ii. The switch is set [-], e.g. in the case of the pro-drop parameter the switch is set to non-pro-drop. Children initially assume that INFL governs the subject in all languages and so they need evidence to set it differently in pro-drop languages; children learning English would use one setting from the beginning and would have no need to change it; children learning Italian would start with a non-pro-

drop setting and would change with time, triggered by evidence (also called parameter resetting).

- iii. The switch is set to [+], e.g. in the case of the pro-drop parameter, the switch is set in the reverse position, i.e. pro-drop; those learning non-pro-drop languages are now the ones who require evidence; those learning Italian need no extra evidence, those learning English do.

2.3.6.1 *Very Early Parameter Setting*

Very Early Parameter Setting (VEPS), proposed by Wexler (1998), claims that children set correctly the basic parameters at the earliest observable stages, that is, at least from the time that the child enters the two-word stage around 18 months of age.

Wexler (ibid.) argued that certain aspects of morphosyntactic inflectional development emerge later in children's development, regardless of the early parameter setting. This unfolding of inflectional development may differ in different languages (Wexler, ibid.): 'the learned parameters (inflectional properties) interacting with the innately unfolding aspects of inflection combine to create quite different surface effects in the development of different languages' (Wexler, 1998).

2.3.6.2 *Very Early Knowledge of Inflection*

According to Very Early Knowledge of Inflection (VEKI⁶) children know the grammatical and phonological properties of inflection components in the language they learn at the earliest observable stage, i.e. from the time that the child enters the two-word stage around 18 months of age (Wexler, 1998).

⁶ VEPS and VEKI come close to being equivalent.

Central to the present study, one set of inflectional items known to very young children (around 18 months of age) are the agreement morphemes (Wexler, 1992, 1994).

2.3.7 The minimalist perspective

In Chomsky's minimalist programme (1992, 1995, 1998, 1999) the notion of LAD and UG is very different to the older versions discussed above (cf. P&P theory). The main distinction proposed is between competence and performance, and therefore, language consists of two components, i.e. the lexicon and the computational system (Chomsky, *ibid.*). The lexicon is where the lexical items are stored with their features, i.e. syntactic, semantic and phonological; lexical items are selected from the lexicon and a number of operations from the computational system is performed (e.g. Move, Merge) to generate a linguistic expression (Chomsky, *ibid.*). The representations of the resulting expressions are mapped onto two layers namely, Logical Form (LF) for semantic interpretation, and Phonetic Form (PF) for phonetic realisation. Consequently, the representation splits up at the 'spell out' stage and according to the Principle of Full Interpretation (PFI) the PF and LF layers may only contain features necessary for the interpretation i.e. a pair of sound and meaning (Eisenbeiss, 2009). Therefore, the process of language acquisition is a 'function mapping experience to a language (Chomsky, 1995:169).

In such a minimalist approach, UG is the set of rules that define all languages and children's initial state S_0 is subject to UG (Chomsky, *ibid.*). The modified P&P model includes a set of principles that apply to all languages and a set of parameters whose values are subject to linguistic variation.

2.3.7.1 *Features*

As mentioned in the previous section, lexical items are stored with their individual set of features (feature bundles). Such features are of three types, i.e. semantic, phonological and formal (Chomsky, 1995). Semantic and phonological features are necessary for interpretation of the LF and PF levels respectively. Formal features are part of the inherent characteristics of lexical items, i.e. they are embedded into the lexical entry. Examples of formal features are category features [\pm nominal], person [3rd], number [\pm plural]–[\pm singular] etc., and, operator features [wh], [Tense].

Features are also divided with respect to their interpretability at PF or LF levels (Chomsky and Collins, 2001). In a given language, a feature is interpretable if it is realised morphologically or phonetically, therefore PF is related to phonetically realised features whereas LF is related to semantically represented features; the latter also involves formal features discussed above. Operator features (syntactic) must be eliminated until the spell-out stage as they are uninterpretable semantically or phonologically, but they play a significant role in syntax (Eisenbeiss, 2009). Features can be interpretable or not depending on the language (Adger, 2003).

2.3.7.2 *Defining feature bundles for MG related to the current study*

Features always have values, and these values are usually binary [\pm] i.e. [+] or [-], with the exception of [modal particle] and [aspect] where more values are available e.g. a verb can be [+perfective, -imperfective], [-perfective, +imperfective] or [-perfective, -imperfective]⁷.

⁷ Single stem verbs are verbs without morphological aspect distinction.

According to Adger (2003), the choice about which feature system is right is purely an empirical one by looking at the facts of a given languages. Features are motivated based on the fact that i) they have an effect on the shape of words (morpho-phonological features), and ii) they have an effect on semantic interpretation; if they have no effect on either the shape or the semantic interpretation, then these features express a syntactic relation (Adger, *ibid.*).

Based on all the above on features, Table 2.1 below, presents a summary of the analyses related to the features discussed in the following chapters:

Table 2.1: Bundles of analyses related to the features of the verbal domain in MG

| Phenomenon | Category | Analyses | Chapter |
|-------------------|--------------------------|---|----------------|
| Tense | Past | regular sigmatic augment saliency | §4 - (§5) |
| Inflection | Finiteness Inflection | modal particles aspect telicity | §5 |
| Agreement | Number Person | singular plural 1 st , 2 nd , 3 rd 3-suffix | §5 - §6 |

Table 2.1 presents the bundles of analyses related to the features examined in this thesis divided into category to reflect the phenomenon analysed in the following sections.

2.3.8 Maturation and Continuity accounts of Language Acquisition

Given the theory of UG, the nature of language development has created great controversy. The obvious question that arises from the claims of Universal Grammar and the Principle and Parameters model discussed above is the following:

If the basis of the language faculty is innate and learning is a matter of switching the parameters to the right position, then why is language acquisition not instantaneous and why is children's language often full of errors?

To account for the above question, linguistic research proposed two accounts, namely the Continuity account (Clahsen, 1991), and the Maturation account (Felix, 1991), discussed in the following sections.

A variety of positions have been advocated, often within either the Continuity account (Pinker 1984, Poeppel and Wexler 1993, Wexler 1998) or the Maturation account (Borer and Wexler 1987, Radford 1988, 1990, Rizzi 1994, Wexler 1994, among many others). Both Continuity and Maturation views come in a variety of formulations.

2.3.8.1 *Continuity*

The continuity hypothesis claims that all principles and parameters are available to children from birth, being part of their innate endowment, but they are not available to be used from the beginning until the child is able to identify and process the words and sentences associated with this specific principle/parameter (Clahsen, 1991).

There are at least two flavours of the Continuity Hypothesis, often referred to as ‘strong’ and ‘weak’ continuity, respectively. Strong continuity accounts claim that all UG properties and principles are available to the child from the initial state. The differences between child language and the adult target are explained through UG-external factors, e.g. constraints on phonological production (Demuth, 1996; Gerken, 1994), interface problems between different linguistic components (Phillips, 1996), or processing limitations (Poeppel and Wexler, 1993). However, it is not clear why some of these factors are qualified as UG-external.

The weaker account of the Continuity hypothesis maintains that although properties and principles do not mature, they may be present in an underspecified version when compared to the target language. Children’s language may be different from the adult target language merely because of parameter/feature setting/mis-setting, e.g. optional stages in children’s grammar are assumed to be due to the underspecification of functional categories (Deprez, 1994; Radford, 1995). According to this weaker version of the Continuity Hypothesis, children’s language is consistent with the adult language since underspecification is a possible setting in adult grammars too.

2.3.8.2 *Maturation*

In its strongest formulation, the Maturation hypothesis claims that children do not have immediate access to all principles throughout development. The assumption is that some principles and parameters are subject to maturation and they become available gradually at certain stages of a child's development (Borer and Wexler, 1987; Felix, 1991).

According to the Maturation Hypothesis, categories and principles are subject to developmental change in children's grammatical knowledge, e.g. functional categories are not present in early child grammar but they mature at a later stage (Radford, 1990; Meisel 1995, Platzack 1992, Tsimpli 1996); this explains the absence of elements such as determiners and auxiliaries in early child speech. During language acquisition certain components of children's knowledge mature and positive evidence from the input allows children to reanalyse those structures (Borer and Wexler, 1987). In this sense principles and parameters of UG are only fully operative at the steady state S_S .

Other hues of the Maturation account suggest instead, that only some principle are subject to maturation while others are fully realised from the initial state (functional categories are not necessarily absent), e.g. a principle maintaining that Tense must be obligatorily projected (Wexler, 1994, 1996), a principle according to which root clauses must necessarily consist of a full Complementiser Phrase (Rizzi, 1993/1994).

2.4 Usage-based accounts of Language Acquisition

Usage-based accounts to language acquisition take a different approach to language acquisition with the theory of emergentism, i.e. the central processing principle is that

language structure emerges from language use. Early language acquisition is a cognitive and social-cognitive activity, rather than a system of broad-based rules.

Usage-based theories hold that the essence of language is its symbolic dimension, with grammar being derivative (Tomasello, 2003, 2006). Children, initially, learn concrete pieces of language, of different shapes and sizes, often only part of larger and more complex constructions (Tomasello, *ibid.*).

According to usage-based theories, UG does not exist and acquisition of language is a single-process in the sense that children learn both regular and irregular constructions in the same way (Tomasello, 2003, 2006). Tomasello (1992) suggests that rules are not the best way to characterise linguistic competence and they are not needed for 2-year-olds.

MacWhinney (1978) recognised three processes central to accounts of language acquisition, namely, rote, combination, and analogy.

According to usage-based theories, children's early one-word utterances are considered to be *holophrases* that often express the same communicative intention as that of the adult expression (Tomasello, 2003, 2006). Multiword utterances, learned and used later, come in three types:

- Word combinations or holophrases are multiple units comprising of concrete pieces of language (but not categories) with a single intonational contour e.g. *ball table, Andrew bed*.
- Pivot schemas are phrases where the structure is defined by one word with another linguistic item filling in a variable slot e.g. *more car, more cereal*. This is a widespread and productive strategy for children.

- Item-based constructions are pivot schemas organised locally but without the presence of syntax (often direct reproduction of what children have heard from adults) e.g. *all done milk, all done now*.

(Tomasello, 2006:114-117)

Pivot schemas are used productively in the sense that children create well defined linguistic categories that correspond to types of linguistic items such as *things I want more*, or *things that are done (finished)* etc. although children do not make generalisations across pivot schemas (Tomasello, 2006). Pivot schemas are therefore constructional islands (cf. Verb-Island Hypothesis) in the sense that children's grammar is not extensive yet (Tomasello, *ibid.*).

According to Tomasello (*ibid.*), later in development (3 years of age) children learn verb classes subject to the following three constraining processes:

- Entrenchment
- Pre-emption
- Knowledge of the semantic subclasses of verbs

Entrenchment is a syntactic constraint referring to the frequency of a verb structure in a child's input; i.e., the more frequently children hear a verb used in a particular construction, the more its usage is entrenched to this construction, thus, disallowing children from extending their knowledge of this construction to a novel verb construction (Tomasello, 2000, 2003, 2006).

Pre-emption refers to the generalisations inferred of the communicative function of a particular linguistic construction (Tomasello, 2000, 2003, 2006). When a child hears a construction that serves the same communicative function as some possible generalisation, the child infers that the generalisation is not conventional, i.e., the construction they heard pre-empts the generalisation (Tomasello, 2000, 2003, 2006): ‘for example, if a child hears *He made the rabbit disappear*, when she might have expected *He disappeared the rabbit*, she may infer that disappear does not occur in a simple transitive construction - since the adult seems to be going to some lengths to avoid using it in this way – the periphrastic causative being a more marked construction’ (Tomasello, 2003: 178).

Knowledge of the semantic subclasses follows closely pre-emption; i.e., when children group together verbs very similar in meaning (Tomasello, *ibid.*). Overtime children make generalisations about which semantic subclasses can be generalised to particular constructions (Tomasello, 2003).

Tomasello (*ibid.*) maintains that often, the constraints of entrenchment and pre-emption may both apply to a verb construction, as a verb that is highly entrenched in one usage is not used in some other linguistic context but an alternative is used instead.

According to Tomasello (2003, 2006), entrenchment is operative from early stages, i.e. before 3;0, as particular verb island constructions become either more or less entrenched depending on usage.

Pre-emption and semantic subclasses begin to work sometime later, perhaps not until 4 years of age or later, as children learn more about the

conventional uses of verbs and about all of the alternative linguistic constructions at their disposal in different communicative circumstances.

(Tomasello, 2006: 52)

2.4.1 Role of input

The role of input in usage-based theories of language acquisition is of major importance. Children learn what they hear from the adults around them; children attempt to learn whole complex adult utterances although sometimes they manage to learn only parts of these often producing only simple words of the original utterance (Tomasello, 2006).

Tomasello (ibid.) maintained that from the beginning children attempt to learn speech acts with a specific communication function corresponding to the adult speech rather than isolated words.

The study of the linguistic expressions and constructions that children hear in their daily lives is the subject of many usage-based studies. For example, Cameron-Faulkner et al. (2003), in their study of child directed speech, examined the speech of English-speaking mothers during their interaction with their 2- to 3-year-old children. The overall findings were that children are exposed to a great number of utterances per day (5,000 to 7,000) although of very fragmentary nature (more than 20% of utterances were not full adult sentences, but instead a noun or a prepositional phrase) (Cameron-Faulkner et al., ibid.). Moreover one-third of the adult utterances were questions and about one-quarter imperatives and copula; only about 15% had the canonical English SVO form (Cameron-Faulkner et al., ibid.).

Tomasello (2006) argued that Cameron-Faulkner et al.'s results demonstrate that child oriented speech is not as complex as written and adult directed speech; therefore

the study of how children acquire language must focus on what children hear rather than presume that children are exposed to the adult language as we know it.

The majority of child directed speech comprises of *highly repetitive item-based frames*, and thus, according to Tomasello (ibid.) the more complex constructions children produce are only a small percentage and mostly based on these highly frequent simple item-based frames.

2.4.1.1 *Trace-back procedures*

Usage-based studies investigating language acquisition (both children's spontaneous data and child directed speech) often use trace-back procedures to establish the connection of children's constructions firstly to their previous speech production and secondly to the adult input. Utterances from a later stage of the child's spontaneous data are traced back in the child's previous production and the adult's speech (Lieven et al., 2003; Lieven et al., 2009).

A matching procedure is employed (performed through programmes such as Autotracer) using deriving operations, e.g. substitution, addition etc. (Lieven et al., 2003; Lieven et al., 2009).

What usage-based studies seek to demonstrate is, that novel and more complex utterances produced by children at a later stage of their linguistic development are based on and can be traced back to their previous production, i.e. entrenchment in combination with some simple deriving operations. The suggestion is that children initially learn rote and consequently build on their knowledge of the distributional values of the frames they already use to create more complex constructions. These phrases are stored in the lexicon and frequency determines what is registered in memory.

The most common operations (substitute and add-on) are the processes used by the child to retrieve old utterances from the memory and build up novel ones.

2.4.2 *Verb Island Hypothesis*⁸

The Verb Island Hypothesis, postulated by Tomasello (1992), claims that the verbs in children's early language are islands. According to Tomasello (ibid.), each island develops its own mini-syntax independently of the other verbs. Furthermore, children learn simple patterns by imitation while the more complex ones develop from the simpler antecedents for each verb separately. Verbal patterns are transferred in an inter-item fashion.

The Verb Island Hypothesis (Tomasello, 1992) also assumes that young children learn grammatical rules for specific words in a piecemeal fashion. Such verb specific knowledge entails the absence of systematicity and associations between different verbs in the child's inflectional system.

Tomasello (ibid.), within the Verb Island Hypothesis framework, proposed two stages in the acquisition of verbal constructions in children's speech: during the first stage, children's verbal knowledge is lexically based attached to particular verbs. Children's grammar is in this respect insular, since each verb has its own internal structure; later, at stage 2, children develop abstract categories and constructions where the verbal system is now interconnected.

Tomasello (ibid.) proposed two different learning processes:

⁸ A detailed analysis of Maria's and Eve's data based on evidence from the Verb Island Hypothesis can be seen in Appendix A.

- a) Children learn simple word combinations around a specific verb by imitation; this is called the pre-grammatical phase.
- b) Later on, children construct more complex structures by adding new arguments to the verbs they already know and use.

Therefore, children's itemised syntax cannot be generalised from one verb to another. Lexical items are insular developing on their own and connections between different verbs are not possible. Learning can only be vertical but not horizontal (Ninio, 2003).

There is a plethora of research papers providing contra-evidence for the Verb Island Hypothesis. Ninio's (2003) findings suggested that children can produce complex structures at the earliest stages of language learning and are able to transfer their knowledge from one item to another, building on less complex sentences they learned first. Combinatory similarities between verbs facilitates such transfer (Ninio, *ibid.*). Ninio (*ibid.*) proposed an 'item-and-analogy' system where children's learning of syntax on a lexical, item-specific basis, improve their learning development by using analogy.

McClure and Pine (2002, cited in Ninio, 2003, p. 6) showed that children after the first use of a construction are able to produce more complex utterances with a greater number of verb arguments than the ones used earlier. The study by Abbot-Smith and Behrens (2006) showed that children learn new constructions based on combinatorial patterns already known to them, thus, children's learning of itemised verb patterns facilitates their learning of similar patterns for other verbs.

2.4.3 Frequency

Usage-based accounts maintain that frequency of occurrence (adult input) plays a vital role in language acquisition.

According to Bybee (1991), children follow two general learning principles:

1. The most often repeated experiences (in production and perception) have the strongest representation and;
2. New experiences are analyzed and stored in terms of existing representations.

Frequency, therefore, is an essential factor, and lexical storage is affected by language use (Bybee, *ibid.*).

Bybee (2007) proposed that memory for language consists of a large store for units of varying sizes with varying degrees of strength, productivity and connection with other units. Bybee (*ibid.*) formalised the importance of frequency with the Network Model, according to which the degree of lexical strength of a word in the lexicon is directly related to this word's token frequency, i.e. words with high token frequency have high lexical strength and therefore are easy to access.

2.4.4 Pre- proto- and morphology proper

Pre-, proto-, and morphology proper are concepts of a constructivist (or usage-based) model of acquisition developed by Dressler, Karpf, Bittner, Kilani-Schoch amongst others. According to Dressler and Karpf (1995), children's morphological development is divided into three stages: pre-morphology, proto-morphology, and morphology

proper. Those three stages present the following characteristics (Dressler, 2003; Bittner et al., 2003):

- A. During the pre-morphological stage, children use only rote-learnt forms. Principles of constructivism and naturalness are the basis of grammatical precursors that will later become grammatical rules. The morphological system is not defined or separated yet from the general cognitive system.
- B. In the proto-morphological phase, children start detecting morphology. Grammatical oppositions emerge which develop into mini-paradigms⁹. During this stage, children construct morphological patterns by analogy and evidence of first inflectional rules may be found in their speech. Generalisations based on early grammatical forms lead to morphological productivity. Morphology becomes a distinct entity, separated from phonology and syntax.
- C. Morphology proper refers to the phase when children advance toward the adult model. During this stage, children create complex constructions and the module of inflection is now dissociated from other morphological modules, e.g. word formation.

The constructivist model applies basic parameters of Natural morphology to the acquisition of inflection. Natural morphology, founded in 1977 by Mayerthaler, Dressler and Wurzel, is the theory of the integration of concepts of markedness and natural phonology (Dressler, 2006). This model does not assume an innate

⁹ A mini-paradigm is a non-isolated set of minimally three phonologically unambiguous and distinct inflectional forms of the same lemma produced spontaneously in contrasting syntactic or situative contexts in the same month of recordings (Bittner, Dressler, and Kilani-Schoch 2003: xxxix).

morphological module but is constructivist, i.e. based on principles of pattern selection and self-organisation: pattern selection means that the child selects some forms in some contexts due to token frequency and salience; self-organisation means that children do not merely imitate input elements but construct themselves their patterns in reaction to the intake (Bittner et al., 2003).

The pre-/proto-morphological approach is compatible with the usage-based approach to language acquisition (Tomasello, 2003) and the network model (Bybee 1991, 1995; Bybee and Hopper, 2001).

Along similar lines moves the network model developed by Bybee (2007), according to which token frequency determines the lexical strength of an item in the lexicon. Higher token frequency of an item increases its lexical strength, and thus, this item has greater lexical autonomy; therefore productivity is determined by token and type frequency (Bybee, *ibid.*).

2.5 Summary

To summarise the theories mentioned above, there are two mainstream theories of language acquisition, i.e. the generative approach according to which children learn language by means of genetic predisposition and the usage-based account according to which children learn and produce what they hear (therefore the more they hear something the easier they acquire it).

The focus of this thesis is on examining the status of elements related to the verbal domain (tense, optionality and inflectional morphology) in the corpus and how the acquisition of these elements develops overtime in children's speech production. The analyses I will be presenting make claims about the linguistic development of past

tense, and inflectional morphology as well as the use of non-adult constructions (root infinitives). My data and findings provide, however, more details of developmental observation compared to previous data analyses.

Some of the predictions and claims I will be making (e.g. high frequency, trace-back, early productivity of morphology) have been previously made by proponents of both generative and usage-based theories. The claims of this thesis, nonetheless, does entail that the findings of the analyses may be compatible with both theories. This suggests that one theory often does not provide a solution to all data under scrutiny. A more immediate question to ask is whether a combination of assumptions made by the two theories can be adequate to account for human language acquisition.

3 METHODOLOGY

This chapter outlines the methodology of this study. It describes how the data were collected, the transcription, coding scheme and the way the analyses of the data were conducted. Methodological problems encountered during the transcription and analyses of the data are also discussed.

The data analysed in this study come from Maria's and Eve's corpus, a new data collection of child Modern Greek spontaneous speech. The information from all data used is as shown in Table 3.1 below:

Table 3.1: Summary of children's general information

| Child | Age of recordings | Sex | Number of recordings | Source |
|--------------|--------------------------|------------|-----------------------------|---------------|
| Maria | 2;0.24 - 2;8.27 | Female | 8 | Doukas corpus |
| Eve | 1;7.15 - 2;11.11 | Female | 13 | Doukas corpus |

3.1 Methods of Data Collection

This section will first provide an overview of different methods of data collection in first languages acquisition, the advantages and disadvantages of each technique and the reasons for the selection of a longitudinal study for this empirical research.

Three types of data can be used for the purposes of investigation and analyses in studies on first language acquisition: experimental data, diary data, and naturalistic data.

Experimental data are the most efficient ones; the data collected by leading an experiment can be extremely useful to a researcher in order to explore specific phenomena of grammatical development and answer questions, as well as to investigate single hypotheses. This is of practical use especially for the examination of infrequent structures occurring in children's speech production. Experimental data can be both cross-sectional, i.e. the experiment is performed to different groups of participants for example different language, and longitudinal, i.e. the same experiment is repeated over a period of time. Although experimental data can be longitudinal, such methodology prevents the researcher from observing the grammatical development and the progress of the acquisition of language because the experiment provides information only for specific times in children's development when the experiment is performed. Other disadvantages of experimental methods are that they can address only one phenomenon in each experiment and they do not show how children produce language in a naturalistic environment. Experimental investigations are appropriate for the study of the different components of language acquisition and for cross-sectional comparison, but, as opposed to naturalistic data, they lack control over the development of language acquisition process of each child across time.

The second type, diary data, is useful for studies in which observation of the early stages of grammatical development is important, where children acquire new grammatical constructions. The advantage is that observation is possible at the very beginning of the language acquisition process; however, when children learn new phenomena in a much faster manner, i.e. the length of utterance increases and so does the speed of learning and fluency, then it is difficult to keep up with the speed of such a process. Because the notes rely only on the tokens of daily observation, as a result it is

hard to record all utterances used by the child; diary notes also often disregard frequencies of the acquired grammatical features. Diary notes are also very useful when investigating phenomena that develop rapidly and would not be captured if the data collection is not very frequent.

Naturalistic data study the language of children produced spontaneously, in a naturalistic environment. Naturalistic data are usually longitudinal, studying the development of each child over an extended period of time. Such studies usually employ regular tape or video recordings over many months.

Apart from the ability to observe the grammatical development of children's spontaneous speech, naturalistic studies, also allow for the study of the analyses of frequencies of the language components used by the children. The big advantage of the study of frequencies is that the researcher can obtain information about structures and elements, compare correct and incorrect instances of word and phrase samples, check the presence of certain grammatical elements and compare and evaluate different time points from the data collection. Finally, frequencies, allow the evaluation of the development of different phenomena at the same time and the effects they may have to each other.

Naturalistic data, however, have some disadvantages too, such as the inability to elicit and investigate specific phenomena of grammar that do not appear in every day spontaneous speech production. Finally, often a tape recording may not provide any data if the child is unwilling to speak during a recording.

The present study has used a longitudinal naturalistic design because the main purpose of the study is to scrutinize the language development of children's spontaneous speech with emphasis on the growth of the verbal domain. The longitudinal

investigation used here provides comparison of different children in different acquisition phases, and highlights the range of individual variation. With the combination of the two methods, i.e. naturalistic and longitudinal, the acquisition of different phenomena can be investigated as well as structures which are acquired gradually or emerge at the same time. This methodology was selected in order to capture the children's language development focusing on the verbal domain.

One of the disadvantages of the longitudinal design of this thesis is that the intervals between the recordings are not consistent and that some of the recordings are very short. Both Maria's and Eve's recordings were planned in advance and a schedule was created for the information of the families. Unfortunately, it was not always possible to follow this schedule.

The analysis of the children's speech from corpora in the public domain that consist of written transcripts, is a complex task because the context is often not given. The researcher cannot identify, for example in a sentence with an action verb, whether or not the action the child refers to has already occurred. The researcher also does not know if the child refers to themselves or someone else. Yet another difficulty in existing corpora is the variability of the conversation topic. The presence or absence of certain verbal forms, e.g. past tense or future tense, can be, solely, the result of a conversation that refers to an activity in the past or future. As a result, it is very difficult to compare grammatical phenomena across children, sessions, databases or languages. Therefore I collected data from two new children in order to be able to assess the context and other peripheral information of the recording firsthand.

3.2 The data

The data analysed in this study come from the Doukas corpus, a new data collection of two Greek speaking children's spontaneous speech collected in Athens, Greece. The data were collected in two different periods, i.e. the first part in 1997 and the second in 2004/2005.

The transcription of the corpus was done using Latin characters partly based on phonemic transcript of the sounds. A transliteration system was devised in order to provide easy reading of the examples provided from the corpora.

The transliteration system mentioned is given below:

Table 3.2: Transliteration system used for the transcription of the data

| Letters and letter combinations | IPA | Transliteration used | Example |
|---------------------------------|------|----------------------|---------------------------|
| A/α | /a/ | a | Alati (= salt) |
| αι | /e/ | e | Etia (= cause) |
| αυ | /af/ | af | Afstiros (= harsh) |
| B/β | /v/ | v | Vazo (= vase) |
| Γ/γ | /ɣ/ | g | Gamos (= wedding) |
| γι | /j/ | ji | Majirevo (= I cook) |
| γε | /j/ | je | Jelao (= I laugh) |
| Δ/δ | /ð/ | d | Deno (= I tie) |
| E/ε | /e/ | e | Ego (= I/me) |
| Eι | /i/ | i | Irini (= peace/Irene) |
| Z/ζ | /z/ | z | Zografizo (= I draw) |
| H/η | /i/ | i | Imera (= day) |
| Θ/θ | /θ/ | th | Theatro (= theatre) |
| I/ι | /i/ | i | Istoria (= history) |
| K/κ | /k/ | k | Kosmos (= world/universe) |

| | | | |
|----------------------|-----------------------|------|---|
| Λ/λ | /l/ | l | Lino (= I untie) |
| M/μ | /m/ | m | Meno (= I stay/live) |
| Mπ | /b/ | b | Beno (= I come in) |
| N/ν | /n/ | n | Neos (= young man) |
| Nτ | /d/ | nt | Kentro (= centre) |
| Ξ/ξ | /ks/ | ks | Ksenos (= foreign) |
| O/o | /o/ | o | Organo (= organ/tool) |
| οι | /i/ | i | Omios (= similar) |
| ου | /w/ | u | Uranos (= sky/Uranus) |
| Π/π | /p/ | p | Pateras (= father) |
| P/ρ | /r/ | r | Riza (= root) |
| Σ/σ/ς | /s/ | s | Sofia (= wisdom) |
| T/τ | /t/ | t | Telos (= end) |
| τσ | /ts/ | ts | Tsanta (= hand/carrier-bag) |
| τζ | /dz/ | tz | Tzaki (= fireplace) |
| Υ/υ | /i/ | i | Ijia (= health) |
| Φ/φ | /f/ | f | Foni (= voice) |
| Χ/χ | /x/ | ch | Cheri (= hand) |
| Ψ/ψ | /ps/ | ps | Psifos (= vote) |
| Ω/ω | /o/ | o | Ora (= hour/time) |
| consonant+iota+vowel | /δj/ /δj/ /px'/ | Cj-V | End j aferon (= interest-ing) D j avazo (= I read) Ep j asa (= I held) |

3.3 Maria's data

Maria's data (Doukas, 1998) were collected in the following way: Maria was tape-recorded starting from the age of 2;0.24 until the age of 2;8.27 for a total of eight

sessions, once a month for 30 to 45 minutes. The number of recordings and age information from Maria's data is shown in the following table:

Table 3.3: Maria's tape recordings' details

| Maria (file) | Age | MLU | Total utterances |
|-------------------------|------------|------------|-------------------------|
| 1 st tape | 2;0 | 3.8 | 84 |
| 2 nd tape | 2;2 | 3.4 | 149 |
| 3 rd tape | 2;3 | 3.8 | 188 |
| 4 th tape | 2;5.4 | 4.2 | 183 |
| 5 th tape | 2;5.24 | 3.6 | 240 |
| 6 th tape | 2;7 | 3.7 | 139 |
| 7 th tape | 2;8.3 | 3.6 | 118 |
| 8 th tape | 2;8.27 | 4 | 160 |
| Total | | | 1261 |

3.3.1 The data collection procedure and transcription

The recordings containing the speech production of Maria were made in her home, in an environment familiar to her. The recordings were made mostly in my presence. When this was not possible, they were made in the presence of well instructed relatives. Finally, the recordings were made under circumstances of playing or reading fairy tales.

As the recordings were in audio-tapes, it was sometimes difficult to know the objects Maria was using. The guidance of the relatives present in the recordings helped me identify and clarify many cases of incomprehensible speech.

All tape recordings were transcribed, according to the transcription format CHAT (MacWhinney, 2000).

3.3.2 Coding

The data extraction, coding, and all the analyses were created manually. Unclear sequences were coded with xxx¹⁰ in the transcriptions and were omitted from further analyses.

All songs occurring in the data have not been analysed, because they are considered to consist of memorised or repeated expressions. On the other hand total or partial imitations of phrases produced by the adults were included in the analyses.

3.4 Eve's data

Eve's data has been collected in a similar way to Maria's data during the period 2004/2005. Eve is also Greek speaking, based in Athens, Greece. Eve was tape-recorded starting from the age of 1;7.15 until the age of 2;11.11 for a total of 16 tape-recordings, with variable intervals (approximately monthly) of variable duration between 30 and 45 minutes. Eve's recordings have been rearranged into 13 files with some recordings merged when the age of recording was too close. The original and

¹⁰ The symbol xxx indicates incomprehensible or unintelligible speech, not treated as a word.

rearranged number of recordings and age information from Eve's dataset is shown in the following table:

Table 3.4: Eve's tape recordings' details

| Original arrangement | Age at recording | Rearranged files | Age at rearranged file | MLU | Total utterances |
|-----------------------------|-------------------------|-------------------------|-------------------------------|------------|-------------------------|
| 1st tape | 1;7.15 | 1 | 1;7 | 2 | 427 |
| 2nd tape | 1;8.7 | 2 | 1;8 | 1.6 | 362 |
| 3rd tape | 1;9.3 | 3 | 1;9 | 2 | 1120 |
| 4th tape | 1;9.17 | | | | |
| 5th tape | 1;10.2 | 4 | 1;10 | 1.9 | 1075 |
| 6th tape | 1;10.30 | | | | |
| 7th tape | 1;11.16 | 5 | 1;11 | 2.3 | 669 |
| 8th tape | 2;0.04 | 6 | 2;0 | 2.4 | 684 |
| 9th tape | 2;2.25 | 7 | 2;2 | 2.5 | 603 |
| 10th tape/A | 2;2.26 | | | | |
| 10th tape/B | 2;3.23 | 8 | 2;3 | 2.3 | 551 |
| 11th tape | 2;5.26 | 9 | 2;5 | 2.8 | 589 |
| 12th tape | 2;6.21 | 10 | 2;6 | 2.5 | 379 |
| 13th tape/A | 2;9.29 | 11/A | 2;9 | 2.8 | 174 |
| 13th tape/B | 2;10.28 | 11/B | 2;10 | 2.7 | 245 |
| 14th tape | 2;11.11 | 12 | 2;11 | 4.4 | 38 |
| Total | | | | | 6916 |

Table 3.4 shows the original and rearranged recordings of Eve's data. Recordings 3 and 4, 5 and 6 and 9 and 10A were merged because of the short interval between the sessions. The original plan also included further recordings beyond the age of 3 years, but due to an unforeseen illness in the family the recordings were stopped at the age of 2;11. More recordings were also planned on a monthly basis, however, the child and her family were not always available to carry out a session at the requested date.

Eve's recordings started at an earlier age in comparison to Maria. From the analyses of Maria's speech it appears that Maria was in a more advanced developmental stage from the onset of the collection i.e. age 2;0. Therefore it was decided to place the onset of the data collection of Eve earlier at the age of 1;7 hoping to be able to capture and observe the earlier stage of development of grammatical phenomena and as a result the developmental transitions of her speech production. This also conforms with Stephany's corpus (available from the CHILDES) where the lower age bound is 1;9

3.4.1 The data collection procedure

Eve's recordings were also mostly realised in my presence, when this was possible, if not, in the presence of relatives under my instructions, usually the parents. Most of the times, the data collection took place at Eve's home.

3.4.2 Transcription

For the transcription of the tape recordings I used the transcription format of CHILDES, Codes of the Human Analysis of Transcripts i.e. CHAT (CHILDES, MacWhinney,

2000). All transcription marks for repetitions, imitations, unintelligible words etc. were done according to the CHAT format (details can be found in the CHAT manual).

Many of the transcripts were checked by and discussed with Theo Marinis and other native speakers of MG.

3.4.3 Coding

3.4.3.1 Morphological coding

Eve's data were coded for verbal forms and verb related elements such as particles, negation, and clitics. The CHILDES/CHAT coding mode was used and the code was created based on specifications of the programme.

The diagram below shows an example of a coded utterance with the verb features identified:

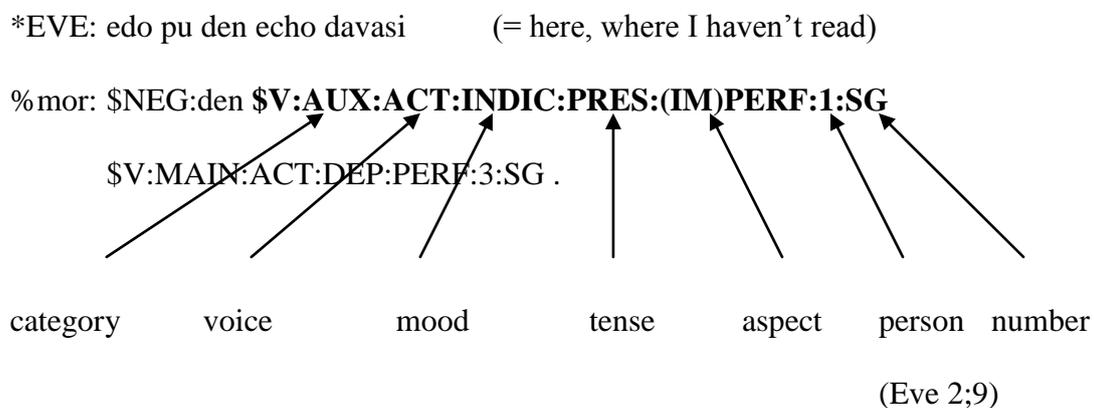


Table 3.5 below shows examples of verbal forms and their coding.

Table 3.5: Examples of verb coding

| Example - Verb | Coding |
|---|---|
| Present Tense Grafi (=he/she writes) Djavazi (=he/she reads) Theli (he/she wants) | \$V:MAIN:ACT:INDIC:PRES:IMPERF:3:SG |
| Imperfect Epeza (= I was playing) Djavaza (=I was reading) | \$V:MAIN:ACT:INDIC:PAST:IMPERF:1:SG. |
| Subjunctive (perfective) na grapso (= to writePERF) | \$PRT:na \$V:MAIN:ACT:SUBJ:DEP:PERF:1:SG |
| Subjunctive (imperfective) na grafo (= to writeIMPERF) | \$PRT:na \$V:MAIN:ACT:SUBJ:DEP:IMPERF:1:SG |
| Simple Past/Aorist Egrapsa (= I wrote) Djavasa (= I read) | \$V:MAIN:ACT:INDIC:PAST:PERF:1:SG |
| Future Tense tha vapso (= I will paint) | \$PRT:tha \$V:MAIN:ACT:INDIC:FUT:PERF:1:SG |
| Imperative with clitic pjas-to (= catch-it) | \$V:MAIN:ACT:IMP:PRES:PERF:2:SG \$CL:en |
| COPULA <i>be</i> Ine (=he/she is) | \$V:COP:ACT:INDIC:PRES:(IM)PERF:3:SG. |
| Clitic + Auxiliary + Main Ta chis vali (= you have put them) | (1)\$CL:pro (2)\$V:AUX:ACT:INDIC:PRES:(IM)PERF:2:SG (3)\$V:MAIN:ACT:DEP:PERF:3:SG . |

Each verb was coded for category, voice, mood, tense, aspect, person and number. To code the corpus, I used the following morphological coder:

| <i>Coder</i> | <i>Abbreviation's key</i> |
|--------------------------------------|---|
| \ +l1 +s1 +b50 | |
| %mor | Morphological coder |
| " \$V \$eV | Correct Verb – Error Verb |
| :MAIN :AUX :COP | <u>Verb category</u> : Main – Auxiliary – Copula |
| :ACT :PASS | <u>Voice</u> : Active – Passive |
| :INDIC :SUB :IMP :UC(Unclear) | <u>Mood</u> : Indicative – Subjunctive – Imperative – Unclear |
| :PRES :PAST :FUT :DEP | <u>Tense</u> : Present – Past – Future – Dependent |
| :PERF :IMPER :(IM)PERF | <u>Aspect</u> : Perfective – Imperfective – (Im)perfective (single stem verbs) |
| :1 :2 :3 | <u>Person</u> : 1 st – 2 nd – 3 rd |
| :SG :PL | <u>Number</u> : Singular – plurals |
| " \$0V | Omitted Verb |

In the coder above, the first distinction is between correct (\$V), incorrect (\$eV) verbal forms, and verb omissions (\$0V). The latter coding applies to incomplete utterances with a verb missing but not to unintelligible utterances marked with xxx.

Verbs are divided in three main categories: main verbs, auxiliaries, and copula *be*. Verbs are coded for voice, i.e. active and passive, and mood, i.e. indicative, subjunctive, imperative and unclear.

In the Mood distinction, the code unclear (UC) was created in order to account for ambiguous utterance when the particle is missing or there is a phonological reduction in the pronunciation, i.e. impossible to identify if *na* is used or *tha* instead.

1. (n/th)a su grafi (= (he/she) [shall/will] be writing to you)
2. (n/th)a su grapsi (= (he/she) [shall/will] write to you)

The example (1) above, with perfective stem, is ambiguous between subjunctive and future because of the phonologically reduced particle¹¹ and (2) with imperfective stem is unclear between subjunctive and future. In both cases, the verb is labelled *unclear* and *dependent* and the stem coding is variable according to the verbal stem i.e. perfective or imperfective.

The code UC (unclear) is also used for cases like the following:

3. *EVE: valo pano (= [to/will] put up)

%mor: \$PRT:0prt \$CL:0cl \$V:MAIN:ACT:UC:DEP:PERF:1:SG.

(Eve 1;7)

¹¹ This phenomenon is been identified in other studies too. Children often use an undifferentiated phonological place-holder *a* instead of the modal particles *tha* and *na* (Stephany 1995, Tsimpli, 2005).

In the above example, the missing particle can be either *na* for a subjunctive construction or *tha* for a future, so the sentence is ambiguous, and therefore *unclear* coding is used.

Subsequently, there is the tense distinction, namely, present, past, future¹² and dependent.

The use of the term dependent (DEP) is based on Holton's et al., (1997) term, i.e. the perfective non-past form of the verb (sometimes called aorist subjunctive). As a result, the abbreviation DEP is used in the coding as the non-past verbal form. This is always used with SUBJUNCTIVE or unclear verbal forms when the particle is missing. This coding was used instead of a non-finite category, in order to differentiate the present/past future tense from the aorist/subjunctive tense *na aplosi*. An example of the code used is given below:

4. *EVE: xxxna to ferume (= we shall bring it)
%mor: \$PRT:na \$CL:pro \$V:MAIN:ACT:SUBJ:DEP:PERF:1:PL.

(Eve 2;2)

The aspect distinction follows, comprising of Perfective, Imperfective and (Im)perfective stem. Aspect can be defined as the way an action is viewed by the speaker at the time of the utterance. The imperfective aspect is used when an action is seen as in progress, habitual, or repeated. The perfective aspect on the other hand is used when an action is seen as a complete whole, or in a neutral way.

¹² Future tense verbs are always coded indicative.

The code (IM)PERF is used for verbs like *kano* and *perimeno*, where the distinction between perfective and imperfective stem is not morphologically marked. The same code is used in imperative verbs such as *ela* (*erchome*), *kita*¹³ (*kitazo-kitakso*) in which case it is impossible to disambiguate the aspect of the stem. (IM)PERF aspect is also used for the copula *be* (*ime/ine*) since there is not morphological aspect marking.

5. *EVE: vate [instead of *vapse* or *vafe*] (= paint[perf/imperf])

%mor: \$V:MAIN:ACT:IMP:PRES:PERF:2:SG

(Eve 1;11)

The above case is a phonological reduction and the question was whether or not to consider this an eV because of the phonology. Here it is impossible to disambiguate between perfective and imperfective stem; the perfective stem was chosen based on the context and also because there are no imperfective imperatives in the corpus (presumably not acquired yet). Additionally, Eve has not yet acquired *ps* clusters in medial positions (there is only a *ps* cluster used word-initially *psarja* in Eve 1;9). Therefore, the verb in the above example was considered not to be an eV.

There is person and number distinction in the coding being the last codes for the verbal morphology.

To code the various elements of the verb's environment I used the following coder:

¹³ The imperative *kita* is ambiguous between *kitakse* (*perfective aspect*) and *kitaze* (*imperfective aspect*) and it's impossible to disambiguate this from the context.

| <i>Coder</i> | <i>Abbreviations' key</i> |
|----------------------|--|
| \$PRT | Particle coder |
| :tha :0tha | <u>Particles</u> : tha – omitted tha |
| :na :0na | na – omitted na |
| :0prt | omitted unidentified particle |
| :a | phonetically reduced particle |
| \$NEG | Negation coder |
| :den :min | <u>Negation</u> : den – min |
| \$CL | Clitic coder |
| :pro :en :0cl | <u>Clitics</u> : proclitic – enclitic – omitted clitic |

The coding for particles includes *na* and *tha* particles and also the omission of each of those, when this can be checked from the context. When the phonological expression of the particle is reduced, then it is impossible to disambiguate between *na* and *tha* and so the unclear coding is used, i.e. *:a*.

Coding for negation includes the two negative markers in Modern Greek, i.e. *den* and *min*. *Min* is used only with subjunctive or imperative mood whereas negation *den* is used with indicative. When *min* is used with subjunctive then the particle *na* is obligatory and through this test the subjunctive mood of the verb can be detected. Both negative particles appear always in pre-verbal position.

3.4.3.2 *Errors and error coding*

All errors have been checked against the adult target, i.e. if the verb or sentence is ungrammatical according to the adult grammar. All utterances that are either syntactically (at the sentence level) or morphologically (at the verb level) ungrammatical according to the adult language were classified as incorrect. The error coder devised is as follows:

| <i>Coder</i> | <i>Abbreviation's key</i> |
|-----------------------|---------------------------|
| \ +l1 +s1 +b50 | |
| %err: | Error coder |
| "\$pho | Phonological error |
| "\$miss part | Missing particle |
| "\$RI | Root Infinitive |
| "\$agr | Agreement Error |
| "\$miss cl | Missing clitic |
| "\$miss v | Missing verb |
| "\$other | Other error |

Errors in the coding are divided in seven categories:

- a. Phonological errors: where the verb is phonologically incorrect (checked from the context) then the verb is coded as eV¹⁴. For example, the sentence below was

¹⁴ For unintelligible verbs marked with xxx the space is left empty in the coding, and not classified as a missing verb, i.e. neither \$OV nor \$eV is used in the coding.

coded as phonological error because the verb used was phonologically alien to the adult verb *skistike*:

6. *EVE: kiki to [instead of *skistike*] (= it was ripped off)

%mor: \$eV:MAIN:PASS:INDIC:PAST:PERF:3:SG.

%err: \$pho.

(Eve 1;7)

- b. Missing particle: where the verb is correct but the utterance is ungrammatical in the adult language because a particle is missing. In this case, the verb is not classified as eV, but the sentence carries the [*] symbol at the place where the missing particle should be placed and then %err coding identifies the type of error. Many of the missing particle errors are involved with a root infinitive error too (see following section for details on Root Infinitives and missing particles) but there are missing particles in verbal forms other than 3rd singular person, which are not Root Infinitives. For example, in the sentence in (7) the subjunctive or future particle (na/tha) is omitted:

7. *EVE: [*] valo pano (= [to/will] put up)

%mor: \$PRT:0prt \$V:MAIN:ACT:UC:DEP:PERF:1:SG.

%err: \$miss part \$miss cl.

(Eve 1;7)

c. Root infinitive: this coding is placed immediately after the missing particle errors because most of the time a RI involves a missing particle. In this case the verb is only in the 3rd singular person, and it is formed with the Perfective stem. In the example (8) below, the dependent verb with perfective stem is a RI in 3rd singular and with no particle:

8. *EVE: i Eva fiji, ena xxx xxx (= Eve [shall/will] leave, one xxx xxx)

%mor: \$PRT:0prt \$V:MAIN:ACT:UC:DEP:PERF:3:SG.

%err: \$miss part \$RI.

(Eve 1;10)

d. Agreement: This type of error is often involved with the two previous ones. Other instances of agreement errors involve the whole utterance, in which case the error is usually checked from the context. When Eve refers to herself in the 3rd singular person this is classified as an agreement error. This is also checked from the context.

9. *EVE: theli tsulit(r)a ta pedja (= the children want[3sg] to go on the slide)

%mor: \$V:MAIN:ACT:INDIC:PRES:IMPERF:3:SG.

%err: \$agr.

(Eve 2;3)

In the example (9) above, the subject is not in number and person agreement with the verb. The entire utterance will be classified as involving an agreement error.

- e. Missing clitic: this error involves the omission of a clitic, the presence of which is vital for the grammaticality of the sentence. In the example (10), in Eve's sentence the clitic before the verb was omitted:

10. *AD1: de ta thelis esi ke ta petas ta pechnidja su?

(= don't you want your toys and you throw them out?)

*EVE: petao

(= I throw [them] out)

%mor: \$CL:pro \$V:MAIN:ACT:INDIC:PRES:IMPERF:1:SG.

%err: \$miss cl.

(Eve 1;10)

- f. Missing verb: when a verb is missing then the sentence is ungrammatical. Incomprehensible utterances marked with xxx are not included in the *missing verb* category. For example, the following sentence was incomplete with the verb omitted:

11. *EVE: ela na to (= come on, let's ...)

%mor: \$V:MAIN:ACT:IMP:PRES:(IM)PERF:2:SG \$PRT:na \$CL:pro \$OV.

%err: \$miss v.

(Eve 2;3)

g. Other: the category *other* contains various errors in the utterance that cannot be classified, usually because it is only a unique case observed in the entire corpus:

12. *EVE: andda, andda, andda <(p)ozaraki chtipise ti Eva to> [*] .

(*lit.*: = the little foot hit Eve)

%mor: \$V:MAIN:ACT:INDIC:PAST:PERF:3:SG.

%err: \$other.

(Eve 1;11)

The example in (12) is labelled as *other* because of the unusual word order which is ungrammatical in the adult language or because it is unclear what the child intended.

13. *EVE: de [*] chalas xxx (= not ruin [it])

%err: \$other \$miss cl.

(Eve 2;5)

Finally, the above example is a one-off use of incorrect negation alongside with a missing clitic error, so this was labelled under the category *other*.

3.4.4 Other remarks on the coding

In the following section some remarks are reported about the coding that have affected the coding process and various decisions on how to code some utterances. These

remarks explain and shed some light to specific fragments from the coding procedure and the utterances involved:

- The verb *echo* (= have) has a dual function: it can be a main verb in main clauses or it can be used as an auxiliary e.g. *echo grapsi* (= I have written). The (IM)PERF aspect code was used for its coding, since the verb is irregular and also aspect distinction is not morphologically marked.

For the coding of participles when used in periphrastic verbal forms the coding was based on Holton et al., (1997) according to whom this is dependent like the 3sg –i form non-finite (homophone) but without a Mood specification (unlike the 3sg –i form). There were only three cases in the Eve’s dataset, where the indicative mood marker was removed manually. In the example below *chtipisi* is a participle, but because of the limited presence¹⁵ of such forms in the corpus it was not necessary to create a further distinction for this component in the coding:

14. *EVE: tora pu che chtipi(si) xxx to ponuse ligo na mu to
chadepthi ke mu perasi.

(= now he/she was hurt, it hurt a little, he/she shall kiss it better for me)

%mor: \$V:AUX:ACT:INDIC:PAST:(IM)PERF:3:SG

\$V:MAIN:ACT:INDIC:DEP:PERF:3:SG \$CL:pro

\$V:MAIN:ACT:INDIC:PAST:PERF:3:SG \$PRT:na \$CL:pro

\$CL:pro \$V:MAIN:ACT:SUBJ:DEP:PERF:3:SG \$PRT:0prt

¹⁵ There are only 2 instances of periphrastic verbs in Maria’s data (both at 2;5,4) and 3 instances in Eve’s speech (one instance at 2;0 and two instances at 2;5).

\$CL:pro \$V:MAIN:ACT:UC:DEP:PERF:3:SG.

(Eve 2;9)

As a result, the coding for participles is as follows:

\$V:MAIN:ACT:DEP:PERF:3:SG

- According to Holton et al., (1997), examples like the following fall under the category of impersonal verbs with a personal pronoun clitic:

15. *EVE: jati de mu aresi to paramithi (na) to davaso.

(= because I don't like the fairy tale, to read it)

%mor: \$NEG:den \$CL:pro \$V:MAIN:ACT:INDIC:PRES:(IM)PERF:3:SG \$PRT:0prt

\$CL:pro \$V:MAIN:ACT:UC:DEP:PERF:1:SG.

(Eve 2;9)

There is not further distinction in the coder for impersonal verbs because there are only two cases in the entire corpus. These were identified manually.

- As mentioned above, (IM)PERF was created for verbs without morphological stem distinction for aspect, e.g. *pao*. Initially, *pijeno*, when used in the contracted form i.e. *pao*, appears to be (IM)PERF and all the instances of this verb were coded with the same label. However, there are many cases where the un-contracted form is used in

which case the stem it is clearly IMPERF. As a result, the entire corpus was re-assessed in order to identify those forms and apply the PERF label in the coding.

16. *EVE: apo do pijenune (= this way we are going)

%mor: \$V:MAIN:ACT:INDIC:PRES:IMPERF:3:PL

(Eve 2;2)

17. *EVE: pigame (as opposed to *pijename*) (= we went)

%mor: \$V:MAIN:ACT:INDIC:PAST:PERF:1:PL.

(Eve 2;11)

Here *pijenune* is clearly imperfective rather than the ambiguous (IM)PERF.

- Verb *thelo* (= want): The problem with verbs like *thelo* is that the perfective stem will be formed as follows: *theliso*. However, this form is very infrequent even amongst adults and moreover the child is very unlikely to know and use the form. Furthermore, *thelo* expresses volition and is a wish/future verb. That is why the perfective¹⁶ form is not commonly used, e.g. adults would say *tha theliso*. As a result, it was assumed that Eve uses the imperfective stem because *thelo* is in the present. The same properties are assumed for the verb *boro-boreso* (= can).

¹⁶ The verb has an inherent future reference when used in present tense so there is no need to use the verb with the perfective stem to express future.

3.4.5 Using the coding

The coding used for the transcription of the corpora was created according to the CHAT manual of CHILDES. Computerised transcriptions and coding allow the researcher to analyse large databases in an automated manner. With the use of codings (morphological or syntactic) computational analyses are available to the researcher. Researchers can examine the development of syntax, phonology, and other grammatical components in language acquisition.

In CHILDES morphological coding tiers are available for a variety of languages such as English, Dutch, German, French etc. but not for Modern Greek. This morphological tier coding was constructed based on guidelines available from CHILDES. The coding was constructed to allow me to use computerised analyses provided by CLAN. This also offers a standard format used in many corpora from different languages.

The main advantage of using CHAT and CLAN is, as mentioned above, the availability of computerised analyses. All previous longitudinal data on Modern Greek (at least the ones available in the public domain) were collected in the 70's and 80's and the codings used are specifically designed to investigate a specific fraction of grammar, e.g. Marinis' coding for Christos was focused on the determiner phrase (DP). Often these codes were not available with the data. Therefore there were no standard codes for Greek. To code my data I had to construct a new coding. The construction and use of such coding, however, has proved to be time consuming and presented the researcher with a series of challenges.

Also, because of this lack of standard codes for Greek, another challenge was to provide an accurate description of the analysed utterances using this code. This part was

particularly difficult since many ambiguities and irregularities are presented in children's speech. During the process of coding there were many times that I had to evaluate and re-think about a sentence and many times the coding had to be changed or adapted.

The aim of this new collection was to have data available for analysis to provide an accurate profile of the children's grammatical and morphological development.

3.4.6 CLAN (*Computerised Language Analyses*)

The CLAN (Computerised Language Analyses) programme was used for the analyses performed for this study. This is a programme designed specifically to analyse data transcribed in the CHAT format.

CLAN can be used to perform a large number of automatic analyses on transcribed data. The analyses include frequency counts, word searches, co-occurrence analyses, MLU counts, interactional analyses, text changes, and morpho-syntactic analysis.

CLAN commands were mostly used through the morphological code tiers in the database. The analysed utterances were coded in the transcription and then these coded lines were used at the analyses for the identification of specific elements of the speech production, i.e. copula be, verbal forms etc. The analyses (programmes) used in this study are the following:

- **FREQ** (frequency): is one of the most powerful programmes in CLAN for frequency analysis. Frequency produces a list of all words used in a file, along with their frequency counts. Frequency analysis was used to calculate the number of

times a word occurs in a file, the number of verbs coded in the morphological tier, the rate of person and number in the verbal paradigm etc.

- **COMBO:** combo uses Boolean search strings to match patterns of letters, words, or groups of words in the data file. Combo offers an insight into all the coded tiers from the data files. Combo was used for the identification of errors through the error coding. Combo analyses were also used to find matches of specific morphological items from the coding such as [+past] tense, clitics, particles etc. When combo finds a match to a search it provides the entire utterance in which the search string matched.
- **MLUw** (Mean Length of Utterance in words): the MLU programme computes the mean length of utterance, which is the ratio of words to utterances. The acquisition of grammatical morphemes (or words) reflects syntactic growth and MLU reflects this growth.
- **MLT** (Mean Length of Utterance in a turn): the MLT programme computes the mean number of utterances in a turn, the mean number of words per utterance, and the mean number of words per turn. MLT was used for the purpose of statistical analyses of the data files, such as the overall rate of coded and analysed words and utterances in the corpus.
- **KWAL:** the Kwal programme outputs utterances that match certain user-specified search words. This programme also allows the view of the context in which any

gives keyword is used. Kwal was used to find and analyse specified pieces from the morphological code tiers such as 3rd persons singular, verbal stem variation (perfective vs. imperfective stem) etc.

3.4.7 Frozen expressions

The role of frozen expressions in children's speech is a very important factor for the study of spontaneous speech. Frozen utterances are unanalysed, formulaic expressions that are not processed and interpreted by the child. The different theories and several conditions proposed for the identification of frozen expressions do not agree with one another. Hickey's (1993) set of conditions, adapted in the current study, is the most comprehensive system for the classification of such occurrences. According to Hickey (*ibid.*), frozen expressions are *unproductive large pieces of unanalysed language*. Hickey proposes nine conditions based on three criteria for their classification, i.e. necessary, typical and graded. Necessary conditions are conditions that cannot be exempted. Grading conditions are not necessary, but if they apply then the judgement of the frozen expression is more secure. Typical conditions are normally applicable but they do not have to be always valid and therefore are subject to exceptions (Hickey, *ibid.*). Hickey's system of conditions is summarised below:

Table 3.6: Conditions for formula identification

| Conditions | Criteria |
|---|---------------------|
| 1. The utterance is at least two morphemes long | Necessary Graded |
| 2. The utterance coheres phonologically | Necessary |
| 3. The individual elements of an utterance are not used concurrently in the same form separately or in other environments | Typical Graded |
| 4. The utterance is grammatically advanced compared to the rest of the child's language (i.e. the grammatical pattern is not represented with different words) | Typical Graded |
| 5. The utterance is a community-wide formula, or one which occurs frequently in the parents' speech | Typical Graded |
| 6. The utterance is an idiosyncratic chunk | Typical Graded |
| 7. The utterance is used repeatedly in the same form | Typical Graded |
| 8. The utterance is situationally dependent | Typical Graded |
| 9. The utterance may be used inappropriately, either syntactically or semantically | Typical Graded |

(Adapted from Hickey 1993, p. 32)

The more conditions an expression fulfils the more stable is its identification as a formula, whereas an utterance which fails some of the conditions might be considered as a dubious case of formulaic expression. Based on the above principle and adopting

Hickey's systems of conditions, Maria's and Eve's corpora were examined for the identification of frozen expressions. This is illustrated in the examples below with the copula *be*:

18. pu (n)to to dedei
 where **is/3sg-it** the dedei (horsey)
 where's the horsey?

(Eve 1;9)

19. pu n-to ato?
 where **is/3sg-it** this?
 Where is this?

(Eve 1;10)

In the examples above, the copula *be* can be considered to be a formulaic expression in the sense that the verb used, attached phonologically to the clitic and often to the wh-element¹⁷, meets at least some of the conditions discussed above:

- The utterance is at least two morphemes long (Condition 1)
- The utterance is phonologically coherent (Condition 2)
- The utterance is a community-wide formula (Condition 5)
- The utterance is used repeatedly in the same form (Condition 7)

¹⁷ Such phonologically merged form is grammatically correct and often used in the adult language.

Copula *be*, even when phonologically merged with the clitic and the *wh*-element is used in both Maria's and Eve's speech in a variety of contexts, suggesting that the examples above are not frozen expressions.

Similar cases were observed with the imperative verbs *oriste*, *efcharisto*, and *ela*. *Oriste* (=order-2pl) a verb in the imperative form, but used as an idiomatic expression used when the hearer wants the speaker to repeat what they have said or sometimes used for emphasis. The English equivalent of such expression will be *I beg your pardon* in a situation where the hearer asks the speaker to repeat the utterance for confirmation. *Oriste* is only used once in Eve's data as an immediate repetition after an adult utterance.

The same happens in the following example, where *efcharisto* (=thank-1sg) is used as an adjective or with reduced verbal function or idiomatic expression in Modern Greek:

20. fhcaristo mama

thank/1sg mummy

Thanks Mummy (lit: I thank you mummy)

(Eve 2;6)

Another verb that was considered to be a frozen expression was the verb form *ela* (=come-2sg). *Ela* is used as an idiomatic expression very often with exhortative meaning to do something (English = *let's or come-on*) and not with the literal meaning of the verb:

21. pantofes xxx ela

sleepers xxx **come/2sg**

come-on (take/bring/give) the sleepers

(Maria 2;0)

Oriste, *efcharisto* and *ela* can be used in the adult speech with their literal meaning or as idiomatic expressions, in both cases grammatically correct. The fact, however, that Maria and Eve use some of these forms only in one of their possible pattern may suggest that these expressions are used in an unproductive and therefore are frozen expressions.

These utterances were checked against Hickey's conditions, in order to identify if the expression accomplishes the requested criteria. Moreover, the context was checked in all utterances with these verbs in order to see if the verb is used with their literal meaning or idiomatically or ultimately with both meanings, e.g. in the case of *ela* Maria and Eve use the verb both with idiomatic and literal meaning to express imperative function; *oriste* and *efcharisto* however are only used with idiomatic meaning this is considered not to be productive, but rather a frozen expression.

The identified frozen expressions (*oriste* and *efcharisto*) were excluded from the analysed data in Maria's and Eve's corpus. However, there were only few instances of them, i.e. 23/71 for Eve and 5/36 for Maria of the verbs discussed above, but less than 1% of all analysed utterance for both Maria and Eve.

3.4.8 *Imitations and self-repetitions*

Imitations and self-repetitions have been identified and they were excluded from all further analyses. This was decided following CHAT standard procedures, according to which exact repetitions or reduced repetitions (exact repetitions of a subpart of the utterance) occur within the last five utterances of the adult's or child's production. The same standardised procedure is adopted in both generative and usage-based studies.

3.5 Stages of development: Age and MLU

The definition of developmental stages is a complicated task. Researchers formulate the stages of language acquisition from the point of view of the language portion they examine.

Language development can be measured either based on the biological age of the children or the mean length of utterance that is index of their grammatical development or vocabulary development. Children present many differences in their speech production and it has been suggested that chronological age may sometimes vary in what regards the emergence and acquisition of various syntactic material and grammatical development.

While age is not a reliable criterion for measuring development because some children acquire grammatical components very early and others later, the mean length of utterance (MLU) has been one of the most widely practiced indices of grammatical development in young children. MLU is calculated in the following way: the total number of utterances in a sample is divided by the total number of words or morphemes. The best known measure, which uses morphemes, was devised by Roger Brown in the 1960s.

According to Brown (1973), MLU is an excellent simple index of grammatical development because almost every new kind of knowledge causes an increase in length: the number of semantic roles expressed in a sentence, the addition of obligatory morphemes, coding modulations of meaning, the addition of negative modalities, and, of course, embedding and coordinating. Brown (ibid.) also found that the linguistic complexity of language unfolded fairly predictably as each stage attained. The MLU is a better predictor of language development than age.

Five stages of development are recognised, based on a division of the length continuum into intervals of 0.5 morphemes. There is a good correlation between MLU and age, but the relationship between MLU and the range of constructions found in a sample is less clear. Predicting the grammatical complexity of a speech sample from the length alone is by no means straightforward, especially as length increases. The presence of new morphemes increases length and this is a sign of the acquisition of new knowledge and of the growth of grammatical development.

Brown's (1973) five stages¹⁸ of development on the basis of the children's MLU, ranged from 1.0 to 4.5. The division of the stage is as follows:

¹⁸ More recent studies in language acquisition identify a sixth stage (or five+) for children's development beyond the 4.5 morphemes.

| <i>Stage</i> | <i>MLU range</i> | <i>Age¹⁹ range (in months)</i> |
|--------------|------------------|---|
| Stage I: | 1 – 2 | 1-26 |
| Stage II: | 2 – 2.5 | 27-30 |
| Stage III: | 2.5 – 3 | 31-34 |
| Stage IV: | 3 – 3.75 | 35-40 |
| Stage V: | 3.75 – 4.5 | 41-46 |
| Stage V+: | 4.5 + | 47+ |

According to Brown (1973), at Stage I, the MLU ranges between 1 and 2 morphemes. Free morphemes are used exclusively. Simple semantic relations are evident e.g. agent + action (Doggie run) or action + object (Throw ball). Language functions such as questions, negation, etc. are often conveyed with intonation.

During Stage II, the MLU ranges between 2 and 2.5 morphemes. Noun and verb phrases are now recognisable units of expression. Some questions markers, inflectional markers and prepositions are also used during this stage.

At Stage III, the MLU ranges between 2.5 and 3 morphemes. Children during Stage III use simple sentences with both nouns and verbs. Questions and negation are used more and in a more adult like format and there is some evidence of the presence of auxiliary verbs.

At Stage IV, the MLU ranges between 3 and 3.75 morphemes. A major milestone is achieved with the use of complex constructions such as embedded and relative clauses.

During Stage V the MLU ranges between 3.75 and 4.5 morphemes. Complex constructions include coordinated sentences and prepositional phrases. The use of

¹⁹ While Brown did not assign ages to the five Stages, others have noted the most typical age range for each stage, given here.

modals and auxiliaries is now more consistent and questions and negation are adult-like. Morphological and syntactic operations are now clearly observed.

From a biological age point of view, Lust (2006) summarises the developmental milestones in infants' linguistic production as follows:

| | | |
|----------------|--------------|--|
| Year 1 | 12 months | First words |
| Year 2 | 14 months | Average about ten words in production (Benedict 1979), often single words utterances. |
| | 15 months | Combinations appear |
| | 17-19 months | Successive single word utterances (Bloom 1973). Beginning of sentence construction. Early language-specific constraint on word order and structure are evident, although utterance length is constrained. Gradual release on length constraint as words begin to be combined into sentences. Early word combinations (Bloom 1973; 1970). Often missing overt inflection, with cross-linguistic difference in how much, and which inflection is missing. |
| 2 Years | | Complex syntax, with various forms of embedding and transformations, becomes evident as early sentences grow in length. Morphosyntax continues to grow. |
| Years | | The essential syntax of a grammar for the language is evident. |
| 3+ | | Certain language-specific properties of grammar, syntax/semantics interactions, and lexicon/syntax interactions continue to develop. |

(Adapted from Lust, 2006, p. 280)

Morphemic MLU, however, presents a few disadvantages that cannot be unmentioned. Firstly, morphemic MLU proves to have complications particularly in languages such as Modern Greek, with rich synthetic inflectional morphology. This is because verb suffixes carry more than one feature, a fact that might be only partly represented in children's early speech. According to Stephany (1981) in highly inflectional languages like MG, it is not always possible to decide on morphemic segmentation at early stages of language acquisition. For the above reasons, the MLU was calculated in words in this study, following the example of previous studies, i.e. Stephany (1981, 1986), Marinis (2003), Alexaki (2003).

3.5.1 *MLU_w*

The current study adapted Brown's criteria from English into Greek using the index with words rather than morphemes, the following figures present Maria's and Eve's MLU_w index:

Figure 3.1: Maria's MLUw

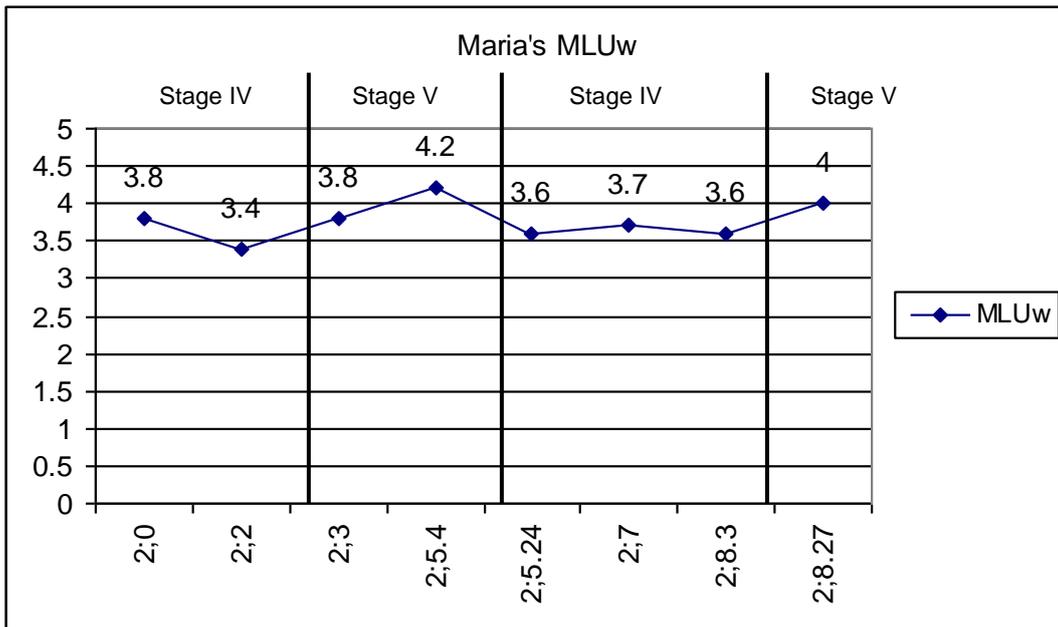
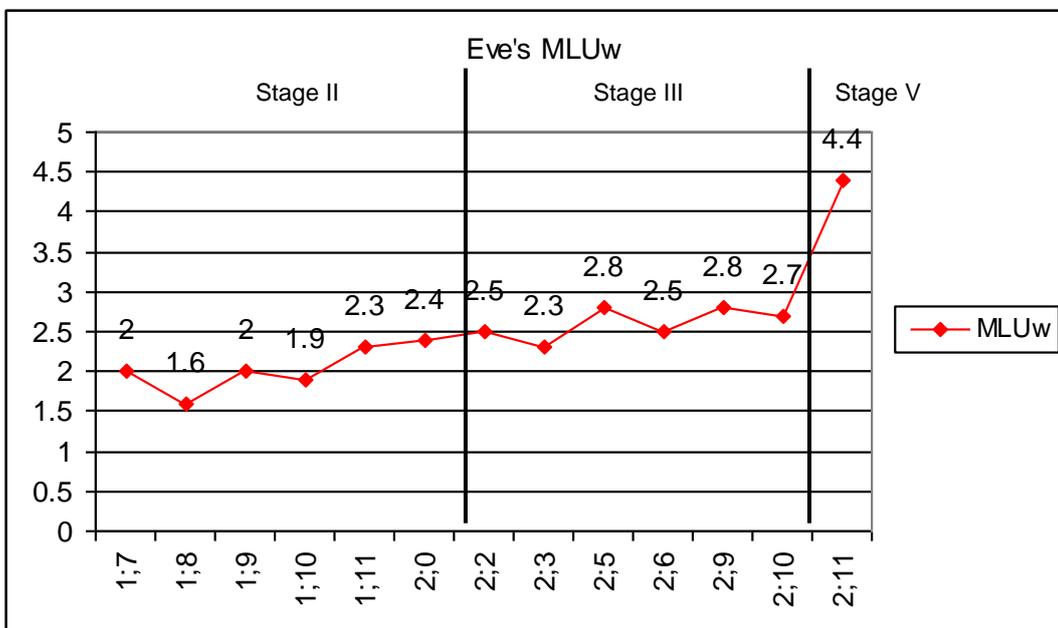


Figure 3.2: Eve's MLUw



Figures 3.1 and 3.2 show the linguistic development of Maria's and Eve's speech production calculated using the MLUw index. Maria's speech production is at Stage IV from the onset of the data collection; then it fluctuates between Stages IV and V throughout the data with minor changes in the numbers. Eve's speech production is at Stage II at the onset of the collection until the age of 2;2. Between 2;2 and after 2;10 Eve's index is at Stage III, and thereafter at Stage V. Detailed references to the various grammatical portions examined in this study and their relation to MLUw are discussed in each chapter.

3.6 Contextual information

Contextual information is important particularly in the way it affects the data and inevitably the analyses (Katis, 1984).

Variations in some of the results presented in the following chapters might potentially be due to the contextual make-up of each recording; for example, in Eve's data at the age of 1;9, the file comprises of two original recordings merged in one and as a result the session is much longer. At the same session, there are also three adults participating during the recording and this signifies greater involvement, prompting and interaction with the child.

At 2;9 Eve spent the day with her grandmother and the recording was not very successful in terms of speech production, since Eve's grandmother finds it difficult to cope with Eve and most of the recording contains some kind of contention.

In Eve's data at the age of 2;5, the contextual situation is of book reading and narration of fairy tales and this might, again, indicate that some grammatical feature is more prominent e.g. the use of reported speech or verbs in aorist.

In Maria's dataset at the age of 2;3 the recording is longer and the adult who performed the session engaged Maria continually with active play and book reading. This might mean that the child uses some grammatical features more than in other recordings, e.g. questions.

On the other hand at 2;5 there are three adults participating during the recording (including the experimenter) and the contextual situation involves a lot of conversation and interactive play. Again, some grammatical feature e.g. present and future tense might be more prominent in this context.

The topic of the discussion in each recording might also affect the presence or absence of certain grammatical features, for example, past tense is the most common tense used in narrative (aorist in MG).

Contextual information has not been taken into consideration in the analyses of the data. However, in the following chapters, there is some discussion of observed findings in relation to the context of the recordings.

4 PAST TENSE

4.1 Aims of the chapter

This study aims to investigate the status of acquisition of the past tense in Modern Greek and attempts to explain the problems raised in previous research through the analysis of a spontaneous speech data collection. The collection of the data started when Maria was 2;0 and Eve 1;7 years old. Therefore the data discussed are produced by children of younger age than the ones analysed in previous studies.

The motivation of this study is to further explore the aforementioned structures and provide more evidence about the acquisition of past tense and offering a more comprehensive analysis of early stage of development.

Two main questions are to be addressed. From a nativist point of view, the first question is related to the [+/-past] feature and the acquisition of it:

Is there a difference between the sigmatic and non-sigmatic past tense in the children's production over time?

The second question originates from a usage-based point of view and considers the trace-back system, firstly concentrating on the child's previous production and consequently progressing to the adult's production:

Can the child's past tense forms be traced back to the child's earlier production and to the adult's past tense forms and which of the past verb forms are novel?

The first question is concerned with the use of the morphological features of past tense, i.e. sigmatic (ruled-based, regular) vs. non-sigmatic (allomorph, irregular) and the mechanism of acquisition of these forms (single vs. dual mechanism). The analyses in the second part of this chapter will aim to answer this question in terms of age and MLU of the children.

Based on the findings of previous studies mentioned in the first part of the chapter, the predictions for this question from a nativist point of view can be summarised as:

- Salience: salient acquired earlier than non-salient forms. As a result:
- Regularity: irregular more salient than regular and therefore acquired earlier
- Augment: +augment more salient than –augment and therefore acquired earlier

The predictions to the second question, from a usage-based point of view focus on the notion of frequency:

- Frequently used past tense forms in children’s speech can be traced back and they are used more in the children’s previous speech production. Frequently used past tense forms in children’s speech are also frequently used in the adult language. Non-frequent past tense forms are often novel and cannot be traced back in children’s previous production. Non-frequent past tense forms are infrequently used in the adult language. The frequency of verbal forms in adults’ language were analysed and determined through the Hellenic National Corpus available from the Institute for Language and Speech Processing²⁰.

²⁰ For more information on the National Hellenic Corpus see section 4.10.2

4.2 Past tense formation in Modern Greek

The Modern Greek (MG) verbal system presents extensive morphological variety with the verb marked for aspect, tense, mood, person, number, and voice. Tense and aspect is the main focus of this study and, in particular past vs. non-past tense and perfective vs. imperfective aspect distinctions.

The first part of this chapter outlines the past tense system in Modern Greek and therefore presents some of the main accounts for the formation of past tense in Modern Greek. Studies on the acquisition of past tense in Modern Greek follow in the second part of the chapter. The final part presents analyses of past tense features from Maria's and Eve's data, the results of the analyses and a short discussion of the findings.

4.3 Aspect

Aspect indicates the difference in the way the action expressed by the verb is viewed by the speaker and presented to the hearer (Holton et al., 1997, 2004). Aspect is a key feature of the MG verbal system. Almost all verbs distinguish between perfective and imperfective verbal stems. Imperfective aspect indicates that the action of the predicate is in progress, repeated or habitual. Perfective aspect indicates completion of the action viewed in its totality as a single and complete event.

MG aspectual information is not lexical, i.e. doesn't come from the verb itself but is instead assigned by aspect morphology. Giannakidou (2002) argues that in MG, unlike English, verbs are unambiguously eventive or stative, depending on whether they have perfective or imperfective aspect. If a verb is inherently stative then perfective morphology should not apply. Thus, verbs with perfective morphology are unambiguously eventives. Although Varlokosta (2002) adopts this view, she also seems

to suggest that aspectual information is lexical in some cases, e.g. state factual/knowledge verbs such as *ksero* ‘I know’. Finally, Giannakidou (2003) proposes that MG verbs contain aspectual information, but it is unspecified. When perfective or imperfective aspectual morphology applies, it becomes specific.

Aspect is marked on the verb stem. The aspectual distinction is, thus, expressed in each verb with the use of two different stems, the *perfective* and the *imperfective*. The perfective stem normally undergoes stem change with the addition of the aspectual marker -s- whereas the imperfective stem remains morphologically unchanged (same as the plain verbal root).

Table 4.1: Perfective-imperfective aspect distinction for verb *graf-o* (= write)

| <i>Aspect</i> | | | |
|---------------------|--------------------|----------------------|--------------------------|
| Imperfective | | Perfective | |
| Present | → graf- | Aorist | → e-graps- |
| Imperfect | → e-graf- | Perfect | → <i>echo</i> graps- |
| Future continuous | → <i>tha</i> graf- | Pluperfect | → <i>icha</i> graps- |
| | | Future instantaneous | → <i>tha</i> graps- |
| | | Future anterior | → <i>tha echo</i> graps- |

Table 4.1 shows the distribution of aspect across the various tenses. Present, Imperfect and Future Continuous are formed using the imperfective stem, whereas Aorist, Perfect, Pluperfect, Future instantaneous, and Future anterior employ the perfective stem.

4.4 Tense (past)

Tense in MG is realised morpho-phonologically through suffixation (*morphologically*) and stress-shift (*phonologically*).

In traditional MG grammars, past is expressed through four different tenses; Imperfect, Aorist²¹ (simple past), Perfect (parakeimenos) and Pluperfect (ypersuntelikos) (Tsopanakis, 1994; Holton et al., 2004).

The Imperfect indicates that the action of the verb was happening in the past in the same way it happens in the present and in future, providing a way to express continuity and repetition. The Imperfect is formed with the affix *-a* in the 1st conjugation, *-ousa* for verbs in the 2nd conjugation, and *-omoun* for passive constructions.

The Aorist tense complements the imperfect, providing information about how aspect, or the way the action happened in the past. In other words, the Aorist tense expresses instantaneity and reinforces the idea of completion of an action in the past. This is in complementary distribution with the continuous action expressed by the imperfect. The Aorist is formed from the root of the present tense and the affix *-sa* for active voice verbs and *-(θ)ika* for verbs of passive voice.

For their formation, both the Aorist and the Imperfect also need the Augment at the beginning of the verbal root which is usually the prefix *e-*. The Augment is essential since the distinction of +/- past is specified in the presence vs. absence of the *e-* prefix in the verbal root. The combination of the past prefix and the aspectual marker results in

²¹ I will use the term *aorist* throughout to indicate the MG simple past being different from other language's simple past, e.g. English.

the two past tense forms known as the Imperfect and the Aorist. In the following example, it can be observed that both the imperfect and the Aorist (as opposed to the present) have the syllabic augment *-e*.

Present: *gráfo* (= I write)

Imperfect: *egráfa* (past-write-imperf-1sg) (= I was writing/used to write)

Aorist: *egrapsa* (past-write-perf-1sg) (= I wrote)

In Modern Greek, augmentation is divided into three types (Holton et al., 1997):

- Syllabic augment – carries the stress in past tense
Gráfo → *egráfa* (= I write → I was writing/used to write)
- Vocalic augment – present in a limited²²/small number of verbs
Elpízo → *ilpíza* (= I hope → I was hoping)
- Internal augment – present in verbs formed with a preposition.
Ipovaló → *ipevalá* (= I submit → I was submitting/I submitted)

The Augment applies only in the active voice of imperfect and aorist of verbs in the first conjugation with a one-syllable stem in singular and 3rd person plural.

Perfect (*parakeimenos*) is a tense with a double identity. Although it belongs to the past, its formation is compound, namely it is formed from the present tense of the auxiliary verb “echo” (= I have) and the non-finite of the aorist of the verb. For this reason, the auxiliary form of the perfect is similar to the present while the non-finite

²² Verbs whereby the initial vowel ε- (e), α- (a) or αι- (ai) of the verb stem is lengthened to η- (i).

aurist form is similar to the past that has already happened. For example, the perfect tense of *grafo* (= I write) is *echo grapsi* (= I have written) and is formed from the present of the auxiliary “echo” (= I have) and the non-finite form “grapsi” (= written) which has a perfective root and remains unchanged. In traditional grammars, *Grapsi* is referred to as the aurist infinitive. The Perfect in MG is closely related to the aurist, in terms of aspect, and the aurist is often used instead of the perfect. Perfect is a telic tense and as such expresses telicity in the past.

According to Klairis and Babiniotis (2004), the speaker using the perfective stem (Aurist, Perfect) perceives the action of the verb in its totality (telic) without taking into consideration the duration or the stages of the action. The speaker using the imperfective stem (Present, Imperfect) perceives the action while it is developing (atelic). The use of either the telic or the atelic condition does not depend on the actual temporal constitution of the action. Each stem can be used for either instant situations or for continuous situations; the difference between the two lies on the different point of view of the speaker.

The Pluperfect (*ypersuntelikos*) is a past tense, formed with the imperfect of the auxiliary “echo” (= I have) and the non-finite form of the aurist (infinitive).

Pluperfect: *icha grapsi* (= I had written)

“icha” (= I had) is the imperfect of the auxiliary verb “echo” (= I have), and “grapsi” (= written) is the non-finite form that remains unchanged in the perfect. The Pluperfect indicates that one action happened before another action took place and the latter action

4.5.1 Analysis 1

According to Holton et al. (1997), Triantafillidis (2002), Klairis and Babiniotis (2004), Clahsen and Stavrakaki (2004), Stavrakaki and Clahsen (2009), amongst others, there are two types of aorist forms: the regular and the irregular one. The Aorist is formed by the combination of the perfective and past categories (Holton et al., 1997). Also, the past tense perfective distinguishes between sigmatic and non-sigmatic forms, depending on the stem of the verb (Holton et al., 1997; Triantafillidis, 2002; Klairis and Babiniotis, 2004; Clahsen and Stavrakaki, 2004; Stavrakaki and Clahsen, 2009).

Furthermore, Klairis and Babiniotis (2004) distinguish three different types of sigmatic morphology:

| | | | |
|-------|---|-------------------|-----------------------------|
| /-s/ | → | * <i>li-s-o</i> | (= untie/NonPastPerfective) |
| /-ks/ | → | * <i>ani-ks-o</i> | (= open/NonPastPerfective) |
| /-ps/ | → | * <i>gra-ps-o</i> | (= write/NonPastPerfective) |

(Klairis and Babiniotis, 2004: 510)

A more detailed description of sigmatic and non-sigmatic morphology proposes three categories for the formation of the sigmatic class and three for the non-sigmatic class (Stavrakaki and Clahsen, 2009; Holton et al., 1997; Ralli, 1988):

Sigmatic

a) Consonantal phonological change in the coda of the verb-stem:

- i. *graf-o* I write

ii. *e-grap-s-a* I wrote

b) Phonological change accompanied by a deletion of the stem-final consonant:

i. *lin-o* I untie

ii. *e-li-s-a* I untied

c) Systematic allomorphy of the perfective stem:

i. *mil-o* I speak

ii. *mili-s-a* I spoke (where *mili-* is the perfective stem of *mil-*)

Non-sigmatic

a) Unpredictable (suppletive) perfective past tense forms:

i. *tro-o* I eat

ii. *e-fag-a* I ate

b) Idiosyncratic stem-internal alterations plus the augment e-:

i. *plen-o* I wash

ii. *e-plin-a* I washed, where the perfective stem *plin-* is an allomorph of the imperfective stem *plen-*.

c) Idiosyncratic stem-internal alterations without the augment e-:

i. *konteno* I shorten

ii. *kontin-a* I shortened

The sigmatic form is considered to be the regular, rule-based form of the past. This form is predictable and morphologically transparent (Stathopoulou, 2009; Stathopoulou and Clahsen, 2010). On the other hand, non-sigmatic past tense forms are

morphologically less transparent than the sigmatic ones, as they do not have a perfective past tense affix but instead have idiosyncratic internal stem changes (Stavrakaki and Clahsen, 2009).

The formation of the past tense in MG is thus portrayed as a dual mechanism where the regularity of verb-forms depends upon the presence or absence of the perfective past tense affix *-s-*. The aspectual *-s-* represents the perfective/imperfective aspect distinction on MG verbs, a distinction which is active in both past and non-past forms.

4.5.2 *Analysis 2*

Ralli (2005) proposed a system that does not distinguish between regular and irregular formation, but instead refers to systematic and non-systematic allomorphic variation in the stem formation of MG verbs. In this system, there is a close relation between verb inflectional classification (and the irregularity of stem) and past formation in MG.

Ralli distinguishes four categories of verbs, which range from no allomorphy through to stem substitution. The stress-shift feature is considered to be part of the lexical properties of the *-a* suffix attached to the past tense:

Verb categories

- a. No allomorphy

lín-o / **é-lis-a** / **lís-o** (= *untie / untied / to untie*)

- b. Non-systematic allomorphy

févy-o / **é-fiγ-a** / **fiγ-o** (= *leave / left / to leave*)

- c. Substitution

lé-o / íp-a / p-o (= *say / said / to say*)

d. Systematic allomorphy

poná-o / pónes-a / ponés-o (= *ache / ached / to ache*)

(*Ralli, 2005: 138*)

4.5.3 Analysis 3

Mastropavlou (2006) proposed a more comprehensive analysis of the Greek aorist formation and also takes into account the stress-shift rule. Mastropavlou suggested that Ralli's classification (discussed in the previous subchapter) reflects only formation with respect to perfective stem, but not to past tense. Building on Ralli's classification, Mastropavlou proposed a more detailed classification that, additionally, reflects the stress shift rule and the syllabic augmentation rule, being the main characteristics of the formation of the past in Modern Greek.

Table 4.2 (from Mastropavlou, 2006) showed the classification of past tense formation and irregularity in order of increasing irregularity (or decreasing predictability).

Table 4.2: Past (simple) tense formation and irregularity

| | <i>Regular</i> | <i>Example</i> |
|----|--|---|
| a. | +stress-shift (phonological) [+SS] | γράφ- <i>o</i> → έ-γράφ- <i>s-a</i> (= write) |
| | +aspectual marker (morphological) [+s] | |
| | –stem allomorphy [-V] | |
| | | |
| b. | +stress-shift (phonological) [+SS] | δίν- <i>o</i> → έ-δίν- <i>s-a</i> (= give) |
| | +aspectual marker (morphological) [+s] | |
| | +stem allomorphy [+V] | |
| | | |
| c. | +stress-shift (phonological) [+SS] | μέν- <i>o</i> → έ-μίν- <i>a</i> (= stay/live) |
| | –aspectual marker (morphological) [-s] | |
| | +stem allomorphy [+V] | |
| | | |
| | <i>Irregular</i> | |
| d. | + stress-shift (phonological) [+SS] | τρέ- <i>o</i> → έ-φάγ- <i>a</i> (= eat) |
| | –aspectual marker (morphological) [-s] | |
| | +stem substitution [+stem] | |
| | | |
| e. | –stress-shift (phonological) [-SS] | βέν- <i>o</i> → βίκ- <i>a</i> (= enter/come in) |
| | –aspectual marker (morphological) [-s] | |
| | +stem or suffix allomorphy [+stem] | |

(Mastropavlou, 2006: 55)

Mastropavlou claims that the sigmatic variation employed in the stem of many Greek verbs is an aspectual marker expressing the perfective/imperfective distinction. The aspectual /s/ is not exclusively typical of the past but applies to non-past forms (future formation *tha pe-ks-o*).

Mastropavlou employs the stress-shift feature [+/-SS] to add another dimension to the distinction of past tense formation. The stress-shift rule employed for the past formation is as follows:

Past tense formation rule:

- (i) +stress-shift +augment: verbs with monosyllabic stems
 –augment: verbs with disyllabic stems

- (ii) suffixation: –a/–es/–e/–ame/–ate/–an

(Mastropavlou, 2006: 50)

Mastropavlou also introduces the issue of *salience*. Perceptual salience is defined in terms of Phonological Form (PF) interpretability. Past tense formation in MG is both morphological and phonological marked and this makes the past tense highly interpretable at the PF level. The use of those morphological and phonological features in past formation also defines the degree of salience. According to this analysis, the morphological feature of augment is a central part of the past tense morphology and the defining factor for salience; that is [+augment] is more salient, whereas [-augment] is less salient. Moreover, the irregular past tense is suggested to be more salient whereas regular past formation is less salient.

There is controversy amongst linguists about the status of the augment. Some studies are against the morphemic nature and importance of the augmentation for the past tense formation, as there is evidence that it is mainly a stress carrier rather than a functional prefix and thus considered to be a phonological element necessary for stress-shift (Ralli, 1988). The same studies propose that the main feature of past is the past suffix (Ralli, 2005). Despite its purely prosodic and phonological nature, other studies suggest that the augment still expresses the past and is the main morphological feature for past formation (Drachman and Malikouti-Drachman, 1994).

According to Ralli (2005), the stress-shift element is a property of the past suffix –a, stored along with its lexical entry. Mastropavlou, building on Ralli, argues that the stress-shift, and the addition of the syllabic augment where necessary, are closely related to the past tense and actually express the past as a notion in MG, and it is this feature that mainly distinguishes the past from non-past in MG verbs along with the past tense suffix –a.

Table 4.3 provides a summary of the accounts mentioned above based on the features each one uses:

Table 4.3: Summary of the accounts on MG past tense formation (features)

| <i>Account</i> | <i>Features</i> |
|---|---|
| Holton et al., 1997 Triantafillidis, 2002 Klairis and Babinotis, 2004 Clahsen and Stavrakaki, 2004 Stavrakaki and Clahsen, 2009 | - Sigmatic/non-sigmatic (three sigmatic and three non-sigmatic formations with a variety of morpho-phonological processes) - Regular/irregular |
| Ralli, 1988 and 2005 | - Systematic and non-systematic allomorphy - Rule-based paradigm |
| Mastropavlou, 2006 | - Stress-shift rule - Aspectual marker (sigmatic/non-sigmatic) - Stem allomorphy |

All accounts employ the aspectual feature i.e. sigmatic vs. non sigmatic. Ralli's account however, defines past tense formation in terms of stem allomorphy whereas Mastropavlou also makes use of the phonological feature of stress-shift.

This study will be based on Mastropavlou's classification since it provides the most comprehensive combination of features, and it will be expanded based on the empirical data analysed to account also for non-active past tense.

From the analysis of Eve's and Maria's past tense verbal forms, there is a further verb category which has been identified but not accounted for in Mastropavlou's classification. This includes verbs in passive voice such as *zaliz-ome*, *kouraz-ome* whose features are shown below:

Regular

-stress shift (phonological) [+SS]

+aspectual marker (morphological) [+s]

+stem allomorphy [-V]

e.g. *zalíz-ome* → *zalístik-a* (= I feel dizzy/felt dizzy)

kouráz-ome → *kourástik-a* (= I get tired/got tired)

I propose this further category to be the last type of regular past formation (d) in Mastropavlou's classification (Table 4.3).

Morphological changes at the verbal stem, such as aspectual marking -s- or stem allomorphy cannot solely account for the formation of past tense, as such operations are also used in non-past formation, e.g. *tha graps-o* (= I will write-future). Aspectual marking and stem allomorphy are rather aspectual features. Also, none of these features apply to the formation of the imperfect (imperfective stem). The use of the three-feature-system proposed in Mastropavlou provides a complete solution that can account for both aspectual and tense morpho-phonological changes in the formation of the past.

Following this classification system, the data in this study will be analysed for the following features:

- sigmatic/non-sigmatic²³
- [+/- augment] (stress-shift rule)
- regularity

²³ The sigmatic/non-sigmatic analysis mirrors Mastropavlou's *aspectual marker* feature.

4.6 Acquisition of past tense in Modern Greek

As mentioned above, past tense formation is subject to regular or irregular inflection. There are two main theoretical accounts that explain past tense morphological processing: the single mechanism and the dual mechanism.

The single mechanism claims that both regular and irregular past forms are acquired and processed in the same way through the employment of a single processing mechanism and without using explicit representation of morphological rules (McClelland and Patterson, 2002; Stathopoulou, 2009 among others). McClelland and Patterson (2002) propose a single mechanism for the acquisition of English past tense where children acquire the past tense inflection gradually. The processing of what is acquired integrates both regular and irregular forms.

Supporters of the dual mechanism postulate the existence of two distinct cognitive mechanisms for the acquisition and processing of the two different morphological types, i.e. regular and irregular morphology. According to this proposal, regular past forms are processed by applying the general morphological rule (e.g. *-ed* suffixation in English) to the verb stem. Irregular past tense forms, on the other hand, are stored in the mental lexicon as independent lexical entries (Pinker, 1990/1999; Clahsen, 1999; Pinker and Ullman, 2002 among others).

The prediction of the dual mechanism is that children would rapidly acquire regular past inflection and this rule initially applies to all verbs regardless of their regularity. Irregular past tense forms are learned and stored in the memory on a one-by-one basis.

In English the development of irregular past tense it has been described as a U-shaped curve (Cazden, 1968; Pinker and Ullman, 2002). Such development occurs in three stages (Cazden, 1968). During the first stage children use the regular past tense inflectional suffix *-ed* but in a non-systematic fashion. At stage two children apply the suffix *-ed* rule to regular and, often, irregular verbs (over-regularisations). In the third stage children begin distinguishing between regular and irregular past tense formation. Pinker and Ullman (2002) suggest that over-regularisations with the possibility that children can store unanalysed words, is sufficient to explain U-shaped development of irregulars.

In MG, several recent studies have investigated the formation of the aorist in children's speech. Some of these have looked at comprehension, some have investigated production and others have looked at both of these elements. All these studies have used a cross-sectional experimental design. This is the first study looking at aorist formation using longitudinal data from spontaneous samples in Modern Greek. The studies which have focused on comprehension will be presented first and then the studies on production.

4.6.1 Stathopoulou (2009)

Stathopoulou (2009) in her doctoral dissertation investigates morphological and syntactic phenomena in Greek speaking adolescents with Down's Syndrome (DS). In reference to past tense morphology, her study investigates whether Greek speaking adolescents with DS exhibit specific problems comprehending past tense morphology. The performance of the participants with DS is compared with a group of typically developing children. The aims of this study were firstly to contribute to cross-linguistic

research of atypical populations with genetic disorders, and more specifically individuals with DS. A further aim was to determine whether individuals with DS have a delay in acquiring certain linguistic structures or their linguistic system as characterised by properties different from those present in typically developing children. In other words, the research questions of this study are whether the language deficit in DS is due to a generic cognitive impairment rather than a linguistic deficiency and whether the characteristics of such a deficit are unique to DS individuals or shared with those who have different congenital syndromes, such as Specific Language Impairment and Williams Syndrome.

Stathopoulou investigated four syntactic phenomena in relation to morphological comprehension by Greek Speakers with DS: syntactic binding, relative clauses, Wh-questions, and past tense. In keeping with the focus of the current study, only aspects of the past tense from Stathopoulou's study will be discussed.

This study also took into account studies of other languages on DS language comprehension. Several studies have suggested that English speaking DS participants have problems comprehending the regular past tense, but their abilities concerning the use of irregular past tense morphology are intact (Fowler et al., 1994). Moreover, studies on Williams Syndrome (WS) claim that children with WS present the exact opposite pattern, namely the mental lexicon where irregular past forms are stored and retrieved appears to be impaired (Clahsen and Almazan, 1998). Both observations provide evidence and support for the dual mechanism for the acquisition and processing of the past tense morphology and also for modularity of the brain.

For this experiment, eight individuals with DS were recruited. Their chronological ages were between 12;1 and 18;7 year and the mental ages ranged from

5;0 to 7;6 years. All participants were Greek native monolingual speakers. The control group comprised of sixteen typically developing children, ten of whom had an average chronological age of 5;2 years, two had an average chronological age of 6;2 years and four had an average chronological age of 7;4 years.

The aim of Stathopoulou's experiment was to investigate whether Greek speaking adolescents with DS exhibit any problems with comprehending of past tense. The experiment also set out to test the processes that Greek speaking adolescents with DS employ in judging perfective past tense forms. This was done through the use of an acceptability judgment task supported by pictures in order to test the comprehension of the perfective past tense of both the DS group and the typically developing control group (*Perfective Past Tense Test* in Clahsen and Stavrakaki, 2004). Sigmatic and non-sigmatic forms of existing and novel verbs (rhyming and non-rhyming) were used in the comprehension tests.

The results showed that DS participants performed better on existing sigmatic verbs than on existing non-sigmatic ones. The difference, however, between sigmatic and non-sigmatic past tense forms was not significant. The results showed that there was a strong preference for the sigmatic past tense form over the non-sigmatic one for the novel verbs that rhymed with existing ones. Finally, for novel verbs that did not rhyme with any existing verb, the DS participants also preferred the sigmatic form of the past tense.

The results from typically developing children largely showed that past forms with sigmatic morphology were preferred to those with non-sigmatic past constructions. For existing verbs, the percentages of sigmatic forms were higher than the non-sigmatic

forms. For novel verbs (both rhyming and not) typically developing children showed a strong preference towards providing a sigmatic response.

In summary, both DS (especially participants at a later stage of development) and typically developing children performed in a parallel manner, showing higher scores on the sigmatic past than on the non-sigmatic one. The control group exhibited a more clear preference for novel rhyming and non-rhyming verbs compared to the DS participants, suggesting that rhyming similarities did not facilitate the process of analogy for the DS group. Stathopoulou's interpretation was that the performance differences between the two groups arise from the DS participants relying less on analogical generalisations than typically developing children. Stathopoulou claimed that participants with DS exhibit difficulties with memory operations such as analogy, i.e. creating a link between a novel word with an existing one. The data suggested that DS participants are more impaired in such memory processes.

Finally, higher scores of both the DS and control group in choosing the sigmatic past in novel verbs supports the claim that the sigmatic past tense is rule-based, and thus, the most commonly used pattern in MG. It is also suggested that this result could be due to frequency effects.

Stathopoulou's study focused on DS participants and even if there is a control group for comparison, the ages of both populations are quite advanced in terms of language acquisition (DS 5;0-7;6 mental age). The control group participants also matched the mental age of the DS group (mean chronological age: 5;2, 6;2 and 7;4). The findings of the study are informative, providing support for the default status of the sigmatic past in both DS and typically developing children. The study has also made a new contribution to past morphology in Greek, showing that there is no obvious

morphological impairment in the comprehension of the past tense by DS participants. There is, however, a lack of developmental studies of children with typical development and the study solely focuses on comprehension.

4.6.2 Delidaki and Varlokosta (2003)

Delidaki and Varlokosta (2003) aimed to test the Aspect First Hypothesis by investigating how Greek speaking children comprehend tense morphology and how lexical aspect affects the use of Greek morphology in early acquisition.

The Aspect First Hypothesis claims that children's use of perfective forms is limited to telic verbs whereas atelic verbs are mainly used with imperfective aspect. Such a claim has received support from many production studies in numerous languages. There is, nevertheless, evidence from comprehension studies which raises doubts about the validity of the Aspect First Hypothesis (Weist et al., 1997).

Delidaki and Varlokosta proposed that MG verbs are specified for both lexical and grammatical aspect. Lexical aspect provides information about telicity, whereas grammatical aspect supplies information about the start and end point of the predicate. Grammatical aspect is marked on the verb stem (perfective-imperfective) while lexical aspect is inherent. Grammatical aspect is independent of lexical aspect, and, therefore, any predicate may be marked with perfective or imperfective aspect regardless of whether they are telic or atelic predicates.

The study investigated the comprehension of present, past and future tense in eighteen Greek-speaking monolingual children aged between 2;7 to 4;0 years. The children were divided into two groups: the younger group consisted of nine children with an age range from 2;7-3;1 years and the older group consisted of nine children

whose ages ranged from 3;2-4;0 years. The experimental task consisted of illustrations of an agent performing actions in three different locations representing three different points in time: an initial location for the past, a middle location for the present and a final location for the future. Six events, three telic and three atelic, were acted out by the agent.

The results of the study showed that the past tense was the most problematic category for all of the participants, and this was particularly true for the younger group. Children performed well in their answers for the present tense task, but they chose the same answer (middle location – ongoing action) when asked a past tense question (instead of the initial location – past tense action). Children of the older group provided correct answers to all types of questions (past, present and future). There is a difference in the interpretation of past between the two groups with comprehension of the past tense appearing to have improved with age. In the analyses of grammatical aspect, the results revealed that the older children performed better with the imperfective past tense rather than with the perfective. In terms of lexical aspect, the proportion of telic and atelic verbs the children used when employing the past tense was almost the same (52.5% for telic and 47.5% for atelic). Finally in terms of grammatical aspect, in the perfective the older group performed better with telic predicates than with atelic ones, whereas there was no difference between the two for the younger group. With the imperfective past tense, younger children provided slightly more correct answers with telic predicates whereas older children provided more correct responses with atelic predicates.

Delidaki and Varlokosta's results suggested that older children comprehend the past tense better than younger ones do. Tense is already developed in children's speech at the age of 2;7 (Delidaki and Varlokosta, 2003). All tenses, with the exception of the

past tense, are almost fully mapped by the age of 3;2. Lexical aspect doesn't appear to play an important role in the interpretation of tense. Delidaki and Varlokosta concluded that the verbs' aspectual information does not considerably affect the acquisition of tense and therefore the Greek comprehension data do not support the Aspect First Hypothesis.

4.6.3 *Stavrakaki and Clahsen (2009)*

Stavrakaki and Clahsen's (2009) experimental investigation focused on typically developing children's comprehension and production of the Greek perfective past tense. The purpose of the study was to investigate what processes children use in producing and comprehending past tense and how these processes develop with age.

The experiment examined the production of the perfective past tense with sigmatic and non-sigmatic marking in 154 typically developing Greek speaking children and 35 adults using existing and novel verbs. The same materials were used for an acceptability judgement task. The children's group had an age range from 3;5-8;5 years of age. The materials consisted of fifty verbs: twenty existing verbs, twenty rhyming novel verbs and ten non-rhyming verbs. The existing and novel rhyming verbs were divided into sigmatic and non-sigmatic forms.

In the production task, Stavrakaki and Clahsen reported a striking contrast between the sigmatic and non-sigmatic perfective past tense forms. Children often overgeneralised sigmatic forms to existing verbs and the sigmatic form was the most frequent response provided for novel verbs (both rhyming and non-rhyming). This was not the case with the non-sigmatic past tense which did not see overgeneralisation outside its category, i.e. it was only used in novel rhyming verb responses. In the

acceptability judgement task, children showed higher accuracy with the sigmatic past tense forms than the non-sigmatic ones. The sigmatic perfective past tense was the preferred option for novel verbs (both rhyming and non-rhyming). This was especially the case with the younger children and the accuracy scores increased with age. The adults showed high accuracy for both types of past tense. These results confirm the observations from the production task that overgeneralisations with the sigmatic tend to be broader than with the non-sigmatic one.

The asymmetry between sigmatic and non-sigmatic morphology, especially in regards to overgeneralisation, supports the claim that the sigmatic past tense is the default formation in MG. The findings of Stavrakaki and Clahsen appear consistent with the dual-mechanism account of acquisition and comprehension. Such a mechanism can accommodate the proposal that the sigmatic past tense is rule-based while the non-sigmatic one is stored and retrieved from the mental lexicon. The preference for sigmatic over non-sigmatic and the overgeneralisation contrast can also be accommodated by the dual mechanism account.

The sigmatic and non-sigmatic perfective past tenses have a different frequency distribution. Stavrakaki and Clahsen (*ibid.*) performed a count of the verb lemmas in a large corpus of Greek words (Neurosoft Language Tools) and found that the vast majority of perfective past tense forms are sigmatic rather than non-sigmatic, i.e. 2119 sigmatic vs. 147 non-sigmatic past tense forms. The authors claimed, that from a connectionist point of view, frequency and the phonological and semantic marking association of the sigmatic past tense can explain the prominence of this form beyond morphology. From a developmental point of view, these results showed that acquiring the non-sigmatic past tense takes longer than the sigmatic past tense. The reason for this

may be that non-sigmatic forms have to be learned on an item-by-item basis and then recalled from memory.

Stavrakaki and Clahsen concluded that Greek-speaking children acquire the sigmatic perfective past tense earlier than the non-sigmatic one. Nevertheless, even younger children understand and produce forms using inflectional morphology and their speech does not only comprise of uninflected lexical items.

4.6.4 Mastropavlou (2006)

Mastropavlou proposed to assess the effect of perceptual salience on the acquisition of the past tense in normal and impaired language development. For the purpose of her study, she proposed to evaluate the degrees of perceptual salience of the past tense through examining how the features of [+/-augment] contributes to the degree of salience of past tense formation with [+augment] considered to be of higher salience than [-augment]. In addition, the feature of [+/-regular] was examined for salience in the sense that the irregular past tense is expected to be more salient than the regular one.

According to the dual mechanism model of language acquisition, irregular past forms are stored in the lexicon and retrieved from memory while regular past forms are processed by rule application (e.g. in English by applying the *-ed* suffix to the verb stem). Therefore, the rule applies by default unless an irregular entry is retrieved from memory (Pinker, 1991). Children are expected to acquire regular past tense earlier than irregular ones. The feature of regularity defines the perceptual salience proposed in Mastropavlou.

The participants of the study were ten children with Specific Language Impairment (SLI) with an age range from 4;2-5;9. The children's IQ was within the

range of normal intelligence, however all presented difficulties with morphosyntactic acquisition. A control group was also recruited consisting of ten typically developing children. Their age range matched the chronological age of the SLI group (4;2-6;0 years old). A second control group of ten typically developing children was also recruited matching the linguistic development of the SLI group and aged between 3;0 and 3;7, determined through their Diagnostic Verbal IQ Test scores.

Mastropavlou's aim was to examine a variety of both verbal and nominal grammatical categories. The study consisted of ten experimental tasks and two pre-tests. For the purposes of this study, only the experiments regarding the past tense will be discussed. In her study, a series of activity pictures in pairs were designed. Using the same verb, each pair presented A. an ongoing activity and B. the same activity after having been completed. The selected verbs had a high frequency of use. The verbs were selected to present three conditions of salience (S) based on the following features of [+/-augment] and [+/-regular]:

- | | |
|---|--|
| 1 st condition: Regulars +S: 18 items | e.g. ζωγράφισι–ζωγράφισε (= <i>draw/drew</i>) |
| 2 nd condition: Regulars ++S: 12 items | e.g. γράφι – έγραψε (= <i>write/wrote</i>) |
| 3 rd condition: Irregulars: 10 items | e.g. πίσι – ήπιε (= <i>drink/drank</i>) |

(Mastropavlou, 2006: 74)

The results revealed that SLI children exhibited lower scores than both of the control groups in all of the tested conditions. There were differences among the three tested

conditions. In the 1st condition (+S) all of the groups achieved their lowest percentages while in the 3rd condition all of the groups achieved their highest scores.

There were also significant differences between the groups, with the SLI group deviating from the two control groups and the language age group showing slightly different results from the chronological age group. Finally, all groups performed better with verbs of higher salience i.e. ++S, +Irr. There was a clear difference in performance between responses to regular and irregular verbs in the SLI group. Mastropavlou claimed that this suggests the existence of two separate acquisition mechanisms for the regular and the irregular morphology. This difference is not claimed to be due to the degree of salience since conditions two and three were not far apart in the SLI group. The picture was different for the control groups where salience had a clear effect in the younger group (language age). Summarising, both SLI and language age group presented salience and regularity effects in their production, whereas the chronological age group did not.

Mastropavlou concluded that salience influences normal language acquisition in the sense that there is a developmental cut-off point where children grow out of this stage at around the age of four. The results provided evidence and support for the dual mechanism account, demonstrating that there are two different routes for the acquisition of regular and irregular morphology.

4.7 Summary

All of the above studies have added to our understanding of how the past tense (and tense in general) is acquired and produced by young children in Greek. A major omission of these studies, however, is the lack of a developmental and longitudinal

record of children's production. There is scope to understand more about the acquisition and the development of past tense from a longitudinal point of view using spontaneous speech data collection where children's acquisition patterns can be identified.

Furthermore, many of these studies (except Delidaki and Varlokosta's) examine the comprehension and production of the past tense in children with some kind of cognitive and linguistic deficit (e.g. Down's Syndrome, Williams Syndrome, SLI) and even when there is cross-sectional analysis with typically developing groups, the focus is primarily on the atypically developing participants.

In addition, the age of the participants (especially the ones of the control groups) are often much higher than a developmental study would usually opt for in order to match the mental (and chronological) age of the atypically developing groups (Stavrakaki and Clahsen, 2009; Stathopoulou, 2009). The result is that the findings correspond to more advanced stages of development which provides insufficient coverage of earlier ages.

There are several drawbacks from a spontaneous data study for example, the inability to elicit and investigate specific phenomena of grammar that do not appear in every day spontaneous speech production. There are, however, many phenomena in children's language that can only be studied through naturalistic data, i.e. data where the linguistic production of a child is documented providing the means for the child's linguistic development to be observed and analysed in a longitudinal way. The present study aims to contribute to determining production of past tense in younger typically developing children, leading to understanding more about their knowledge of the past tense, and if their regular/irregular formation is productive in an adult-like manner from a very early stage of development. Finally, this study aims to use a combination of

analyses (and features) from previous research, thus, providing a more comprehensive analysis. Mastropavlou's study, for example, does not directly examine sigmatic and non-sigmatic past tense formation.

4.8 Analyses

In the next section the following analyses will be presented on the data of Maria and Eve:

- Use of sigmatic and non-sigmatic morphology
- Use of [+/- augment]
- Use of [+/- regular]
- Trace-back procedure
- Effects of frequency

4.8.1 *Sigmatic morphology*

To address whether or not sigmatic vs. non-sigmatic forms appear first in production, the following quantitative and qualitative analyses were conducted for each dataset:

Sigmatic vs. non-sigmatic past tense morphology

Following Ralli's classification, the main distinction taken into account in the current analysis was the presence vs. absence of the -s- suffixation in the verbal stem:

chala-o → chala-s-a (= I spoil/spoiled)

vs. vrisk-o → vrik-a (= I find/found)

As Eve's data is entirely transcribed in CHAT format, there's no sigmatic/non-sigmatic distinction in the morphological coding, and therefore the following CLAN command was used instead:

combo +t%mor +s*\$*past* eve.cha

Using this command, all verbs coded with past tense were identified in the corpus. The distinction between sigmatic vs. non-sigmatic past tense morphology was checked manually. COMBO was opted for over a FREQ analysis because it provides information about the utterance where the context could be checked too.

The verbal forms coded DEP (dependent²⁴) were left out of the classification. Past tense coding also included Imperfect (imperfective stem) and so this was not included in the analysis. Copula *be* and verbs with *single-stem*²⁵ were also excluded from this analysis.

Repetitions of the same verbal form were counted in the analysis, if they were not immediately following the adult utterances (immediate imitation) and they were not occurring within 5 lines of each other. If they fall under the above criteria, repetitions

²⁴ According to Holton et al., (1997), Dependent is the perfective non-past form of the verb (sometimes called aorist subjunctive). The abbreviation DEP is used as the independent non-past verbal form to indicate a non-finite category. This assists to distinguish between the present/past future tense for tenses like aorist/subjunctive *na aplosi!*

²⁵ Some verbs do not have morphological stem distinction between imperfective and perfective, so their past formation is ambiguous since it can be either imperfect (paratitikos) or aorist. This characteristic does not, nonetheless, affect this distribution, simply because in any case these verbs do not form a sigmatic past. Those verbs were left out of the classification. Such verbs are: *kano*, (= I make) *echo* (= I have), *ksero* (= I know), and they referred as *single-stem verbs* (Katis, 1984).

were excluded following the criteria of CHAT providing a comprehensive guide²⁶. Finally, only clear verbal forms were considered in this analysis.

In Maria's dataset the past tense and the sigmatic vs. non-sigmatic distinctions were also checked manually following the criteria above.

Sigmatic and non-sigmatic marking was additionally analysed as *tokens*, *types* and *verbs*. *Tokens* are occurrences of all verbal forms and *types* (even if repeated), *types* are all verbal forms (even of the same verb, but not repetitions) and *verbs* refer to all the different verbs encountered.

4.8.1.1 Results

Table 4.4 presents the distribution of verbs across *tokens*, *types*, and different *verbs* used in the aorist form in the data of Maria divided into sigmatic and non-sigmatic forms.

²⁶ Criteria also adopted in usage-based studies e.g. Lieven et al., (2003).

Table 4.4: Maria's sigmatic/non-sigmatic past tense forms

| Maria | | Sigmatic | | | | | | Non-sigmatic | | | | | |
|--------------|------------|-----------------|-----------|--------------|-----------|--------------|-----------|---------------------|-----------|--------------|-----------|--------------|-----------|
| Age | MLU | <i>Token</i> | N | <i>Type</i> | N | <i>Verb</i> | N | <i>Token</i> | N | <i>Type</i> | N | <i>Verb</i> | N |
| 2;0 | 3.8 | 33.3% | 1 | 33.3% | 1 | 33.3% | 1 | 66.7% | 2 | 66.7% | 2 | 66.7% | 2 |
| 2;2 | 3.4 | 71.4% | 5 | 71.4% | 5 | 71.4% | 5 | 28.6% | 2 | 28.6% | 2 | 28.6% | 2 |
| 2;3 | 3.8 | 57.1% | 4 | 57.1% | 4 | 60.0% | 3 | 42.9% | 3 | 42.9% | 3 | 40.0% | 2 |
| 2;5,4 | 4.2 | 47.6% | 10 | 45% | 9 | 47.1% | 8 | 52.4% | 11 | 55% | 11 | 52.9% | 9 |
| 2;5,24 | 3.6 | 50.0% | 6 | 45.5% | 5 | 50.0% | 5 | 50.0% | 6 | 54.5% | 6 | 50.0% | 5 |
| 2;7 | 3.7 | 54.5% | 6 | 54.5% | 6 | 54.5% | 6 | 45.5% | 5 | 45.5% | 5 | 45.5% | 5 |
| 2;8,3 | 3.6 | 50.0% | 2 | 50.0% | 2 | 50.0% | 2 | 50.0% | 2 | 50.0% | 2 | 50.0% | 2 |
| 2;8,27 | 4 | 37.5% | 9 | 56.3% | 9 | 53.8% | 7 | 62.5% | 15 | 43.8% | 7 | 46.2% | 6 |
| Total | | 48.3% | 43 | 51.9% | 41 | 52.9% | 37 | 51.7% | 46 | 48.1% | 38 | 47.1% | 33 |

In Maria's data there are few instances of the past in the first recording (1-2). These results are not very informative. For the tokens in the following recordings, the sigmatic past tense is used more between the ages of 2;2 and 2;3 than in later recordings. There is consequently a slight variation in the use of both the sigmatic and non-sigmatic past from 2;5,4 to 2;8,3. Finally, in the last recording (age 2;8,27) the use of the sigmatic past decreases noticeably. In terms of types, the use of the sigmatic past is more prominent between the ages of 2;2- 2;3. From that point until the last recording there are only minor variations between the rates of use of the sigmatic and non-sigmatic past tense and this low variation is also observed with the use of verbs.

Table 4.5 presents the distribution of verbs across tokens, types, and different verbs used in the aorist form in the data of Eve divided by sigmatic and not sigmatic forms

Table 4.5: Eve's sigmatic/non-sigmatic past tense forms

| Eve | | Sigmatic | | | | | | Non-sigmatic | | | | | |
|--------------|------------|-----------------|------------|--------------|-----------|--------------|-----------|---------------------|------------|--------------|-----------|--------------|-----------|
| Age | MLU | Token | N | Type | N | Verb | N | Token | N | Type | N | Verb | N |
| 1;7 | 2 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 100.0% | 1 | 100.0% | 1 |
| 1;8 | 1.6 | 0.0% | 0 | 0.0% | 0 | 0.0% | 0 | 100.0% | 1 | 100.0% | 1 | 100.0% | 1 |
| 1;9 | 2 | 76.9% | 10 | 80.0% | 4 | 80.0% | 4 | 23.1% | 3 | 20.0% | 1 | 20.0% | 1 |
| 1;10 | 1.9 | 67.7% | 21 | 61.1% | 11 | 46.2% | 6 | 32.3% | 10 | 38.9% | 7 | 53.8% | 7 |
| 1;11 | 2.3 | 64.7% | 11 | 62.5% | 10 | 66.7% | 8 | 35.3% | 6 | 37.5% | 6 | 33.3% | 4 |
| 2;0 | 2.4 | 50% | 10 | 43.8% | 7 | 46.7% | 7 | 50% | 10 | 56.3% | 9 | 53.3% | 8 |
| 2;2 | 2.5 | 42.1% | 8 | 43.8% | 7 | 40.0% | 6 | 57.9% | 11 | 56.3% | 9 | 60.0% | 9 |
| 2;3 | 2.3 | 40% | 6 | 41.7% | 5 | 45.5% | 5 | 60% | 9 | 58.3% | 7 | 54.5% | 6 |
| 2;5 | 2.8 | 35.7% | 15 | 44.4% | 12 | 47.8% | 11 | 64.3% | 27 | 55.6% | 15 | 52.2% | 12 |
| 2;6 | 2.5 | 31.6% | 12 | 34.4% | 11 | 32.0% | 8 | 68.4% | 26 | 65.6% | 21 | 68.0% | 17 |
| 2;9 | 2.8 | 33.3% | 1 | 25.0% | 1 | 25.0% | 1 | 66.7% | 2 | 75.0% | 3 | 75.0% | 3 |
| 2;10 | 2.7 | 66.7% | 4 | 60.0% | 3 | 60.0% | 3 | 33.3% | 2 | 40.0% | 2 | 40.0% | 2 |
| 2;11 | 4.4 | 66.7% | 2 | 66.7% | 2 | 66.7% | 2 | 33.3% | 1 | 33.3% | 1 | 33.3% | 1 |
| Total | | 48.1% | 100 | 46.8% | 73 | 45.9% | 61 | 51.9% | 108 | 53.2% | 83 | 54.1% | 72 |

In the first two recordings of Eve, very few verbs were used in the aorist (0-1). Therefore, they are not informative. In the following recordings the picture is complex but informative. In terms of tokens, between the ages of 1;9 and 1;11 Eve uses more sigmatic than non sigmatic verb forms. Then, at the age of 2;0 sigmatic and non-sigmatic past forms are used equally. Consequently, between the ages of 2;2 and 2;9 the non-sigmatic past is used more than the sigmatic one. Finally, in the last two recordings (age 2;10 and 2;11) Eve's use of the sigmatic past increases again. In terms of types, the pattern of use is very similar between the ages of 1;9 and 1;11 with the sigmatic past being used more than the non-sigmatic. This pattern however changes after the age of 2;0 when Eve uses more non-sigmatic than sigmatic as the rates of sigmatic gradually fall. Finally, after the age of 2;10 the use of the sigmatic past increases again. For the verbs, there is a sharp reduction in past tense verbs after the age of 1;9 (from 80% to 46.2%) and then an increase at the age of 1;11. Subsequently, the rate of past tense use gradually drops, following the pattern observed for types and tokens, till the age of 2;9 where the use of sigmatic verbs rises once again. These results and what they reveal will be discussed further in the following section.

4.8.2 Productivity of verb inventory

Eve's and Maria's choice of verbs with the sigmatic and non-sigmatic past tense is revealing, in terms of productivity.

In Eve's data, (Table 4.4) in the first recordings until 2;0 the non-sigmatic past tense is mostly used with the same verbs repeated across the corpus, e.g.

vrisk-o → vrik-a (= I find/found)

pern-o → pir-a (= I take/took)

There is of course some variation and use of novel verbs²⁷. However, this is limited compared to the sigmatic use with the sigmatic past tense used more frequently with novel verbs:

chala-o → chala-s-a (= I spoil/spoiled)

kov-o → e-kops-a (= I cut/cut)

rota-o → roti-s-a (= I ask/asked)

pona-i → pone-s-e (= it aches/ached)

This of course doesn't mean that there are no repetitions of sigmatic forms across the corpus e.g. peft-i → e-pe-s-e (= it falls/fell).

In the second part of Eve's data (2;0 onwards) the opposite pattern is observed, i.e. the non-sigmatic past is more used than the sigmatic one. The occurrences of the sigmatic past diminish and the non-sigmatic past is used more frequently and with novel verbs:

plen-o → eplin-a (= I wash/washed)

men-i → emin-e (= it remain/remained)

din-ome → dithik-a (= I get dressed/got dressed)

kriv-ome → kriftik-a (= I hide/hid)

In Maria's data (Table 4.5), there is a slight increase of the sigmatic past forms at the first recordings until 2;3. After that, it appears that sigmatic and non-sigmatic past forms are used in balance.

²⁷ The terms *novel* is used here for verbs that are not used in previous recordings. The definition of *novelty* will become clearer in the trace-back analysis.

4.8.3 MLU_w

An additional calculation was performed for Eve's sigmatic/non-sigmatic past across the MLU stages. Based on Brown's (1973) MLU²⁸ index of linguistic development, Maria's data goes through 2 stages i.e. Stage IV and Stage V, which alternate over time, as shown below:

Stage IV: 2;0 – 2;2

Stage V: 2;3 – 2;5,4

Stage IV: 2;5,24 – 2;8,3

Stage V: 2;8,27

Table 4.6 below illustrates Maria's sigmatic and non-sigmatic distribution across the MLU stages of development.

Table 4.6: Maria's sigmatic/non-sigmatic distribution by MLU_w

| Maria | Sigmatic | | | | | | Non-sigmatic | | | | | |
|-----------|--------------|----|-------------|----|-------------|----|--------------|----|-------------|----|-------------|----|
| | <i>Token</i> | N | <i>Type</i> | N | <i>Verb</i> | N | <i>Token</i> | N | <i>Type</i> | N | <i>Verb</i> | N |
| IV | 60% | 6 | 60% | 6 | 60% | 6 | 40% | 4 | 40% | 4 | 40% | 4 |
| V | 50% | 14 | 48.1% | 13 | 50% | 11 | 50% | 14 | 51.8% | 14 | 50% | 11 |
| IV | 51.8% | 14 | 50% | 13 | 52% | 13 | 48.1% | 13 | 50% | 13 | 48% | 12 |
| V | 37.5% | 9 | 56.2% | 9 | 53.8% | 7 | 62.5% | 15 | 43.7% | 7 | 46.1% | 6 |

²⁸ For a more detailed discussion about MLU see Methodology chapter.

Table 4.6, shows Maria' use of the sigmatic and non-sigmatic by MLU stages. It is difficult to establish a pattern of this use since Maria's data alternates between Stage IV and Stage V over time. For the tokens, sigmatic use is higher during Stage IV than non-sigmatic use, especially between 2;0 and 2;2. During Stage IV between 2;5,24 and 2;8,3 sigmatic is slightly more than non-sigmatic. During Stage V between 2;3 and 2;5,4 sigmatic and non-sigmatic past are equally distributed. For the types and verbs the use of sigmatic and non-sigmatic follows the patterns observed for the tokens. Finally, during Stage V at 2;8,27 sigmatic past is considerable lower than non-sigmatic for the tokens, more prominent however, for types and verbs.

Eve's data goes through 3 stages of development:

Stage II: 1;7 – 2;0

Stage III: 2;2 – 2;10

Stage V: 2;11

Table 4.7 below illustrates Eve's sigmatic and non-sigmatic distribution across the three MLU stages of development.

Table 4.7: Eve's sigmatic/non-sigmatic distribution by MLUw

| Eve | Sigmatic | | | | | | Non-sigmatic | | | | | |
|-----|-----------|-------|-------|------|-------|------|--------------|-------|-------|------|-------|------|
| | MLU Stage | Token | N | Type | N | Verb | N | Token | N | Type | N | Verb |
| II | 63.4% | 52 | 56.1% | 32 | 53.2% | 25 | 36.6% | 30 | 43.9% | 25 | 46.8% | 22 |
| III | 37.4% | 46 | 40.6% | 39 | 41.0% | 34 | 62.6% | 77 | 59.4% | 57 | 59.0% | 49 |
| V | 66.7% | 2 | 66.7% | 2 | 66.7% | 2 | 33.3% | 1 | 33.3% | 1 | 33.3% | 1 |

In Table 4.7, the use of the sigmatic and non-sigmatic past is shown by MLU stages. For the tokens, sigmatic use is higher during Stage II than non-sigmatic use. During Stage III the non-sigmatic past is more prominent than the sigmatic and in Stage V, sigmatic decreases yet again. There are, however, very few verbs which were produced in Stage V because the recording was very short. In terms of types and verbs, the use of sigmatic and non-sigmatic follow the same patterns observed for tokens.

4.8.4 *Augmentation and Regularity*

As suggested in Stavrakaki and Clahsen (2009), Holton et al. (1997), and Ralli (1988) augment is a fundamental morpho-phonological characteristic of past tense formation that applies to both sigmatic and non-sigmatic verbs. Discussed in the first part of this section, the distinction of +/- past is specified in the presence or absence of the augment (-e) in the verbal root.

An additional analysis was thus performed in the morpho-phonological salience as outlined by Mastropavlou (2006) with the sigmatic and non-sigmatic past tense forms checked for the following features:

[+/- regular]

[+/- augment]

The prediction for the above features, are repeated below for convenience:

- Regularity: irregular more salient than regular and therefore acquired earlier
- Augment: +augment more salient than –augment and therefore acquired earlier

In the data presented in the following tables (Tables, 4.8, 4.9, 4.10 and 4.11) the copula *be* and verbs with *single-stem* were excluded from the calculation. Also, all distributions were performed on tokens rather than verbs or types. This is because augment doesn't apply to all persons of a given verb's paradigm (only singular and 3rd plural persons) as mentioned at the beginning of this section. As a result, the analysis of augmentation sought to investigate each past form on an individual basis. In terms of regularity, this is part of the verb's blueprint, rather than due to morpho-phonological modification, and thus, one could have analysed verbs rather than tokens, but for greater consistency and comparison purposes the analysis was performed in tokens.

Starting with the [+/- augment] feature, Tables 4.8 and 4.9 below illustrate the distribution of the past tense with/without augmentation in the speech of Maria and Eve:

Table 4.8: Maria's [+/- augment] in past tense

| Maria/ Age | Sigmatic | | | | Non-sigmatic | | | |
|---------------|--------------|-----------|--------------|-----------|--------------|----------|--------------|-----------|
| | [+augm] | N | [-augm] | N | [+augm] | N | [-augm] | N |
| 2;0 | 100% | 1 | 0% | 0 | 50% | 1 | 50% | 1 |
| 2;2 | 0% | 0 | 100% | 5 | 0% | 0 | 100% | 2 |
| 2;3 | 25% | 1 | 75% | 3 | 0% | 0 | 100% | 3 |
| 2;5,4 | 50% | 5 | 50% | 5 | 27.2% | 3 | 72.8% | 8 |
| 2;5,24 | 33.3% | 2 | 66.7% | 4 | 0% | 0 | 100% | 6 |
| 2;7 | 33.3% | 2 | 66.7% | 4 | 60% | 3 | 40% | 2 |
| 2;8,3 | 0% | 0 | 100% | 2 | 50% | 1 | 50% | 1 |
| 2;8,27 | 55.5% | 5 | 44.5% | 4 | 6.6% | 1 | 93.4% | 14 |
| Total | 37.2% | 16 | 62.8% | 27 | 19.6% | 9 | 80.4% | 37 |

In the first recording (age 2;0), Maria's data show only one occurrence of sigmatic past tense, and thus this is not informative. From 2;2 years of age Maria's use of the sigmatic past tense without the augment is more prominent except at the age of 2;5,4 where the proportion of augmented and non-augmented sigmatic past tense is 50%-50% and the final recording at 2;8,27 years of age where the use of augmented sigmatic past is slightly higher than the non-augmented. The first three recordings present a very low in number of uses of the non-sigmatic past tense. At 2;5,4 years of age, the non-augmented, non-sigmatic past tense is used more often. Between 2;7 and 2;8,3 there are few occurrences of the non-sigmatic past tense, but the use of the non-

sigmatic with augment increases. Finally, at the age of 2;8,27, the non-sigmatic past without augment is the most common usage of past tense.

Table 4.9: Eve's [+/- augment] in the past tense

| Eve/ Age | Sigmatic | | | | Non-Sigmatic | | | |
|--------------|------------|-----------|------------|-----------|--------------|-----------|--------------|-----------|
| | [+augm] | N | [-augm] | N | [+augm] | N | [-augm] | N |
| 1;7 | 0% | 0 | 0% | 0 | 0% | 0 | 0% | 0 |
| 1;8 | 0% | 0 | 0% | 0 | 0% | 0 | 100% | 1 |
| 1;9 | 80% | 8 | 20% | 2 | 0% | 0 | 100% | 3 |
| 1;10 | 52.3% | 11 | 47.7% | 10 | 30% | 3 | 70% | 7 |
| 1;11 | 36.3% | 4 | 63.7% | 7 | 50% | 3 | 50% | 3 |
| 2;0 | 70% | 7 | 30% | 3 | 20% | 2 | 80% | 8 |
| 2;2 | 25% | 2 | 75% | 6 | 36.3% | 4 | 63.7% | 7 |
| 2;3 | 16.6% | 1 | 83.4% | 5 | 33.3% | 3 | 66.7% | 6 |
| 2;5 | 33.3% | 5 | 66.7% | 10 | 55.5% | 15 | 44.5% | 12 |
| 2;6 | 41.6% | 5 | 58.4% | 7 | 15.3% | 4 | 84.7% | 22 |
| 2;9 | 0% | 0 | 100% | 1 | 0% | 0 | 100% | 2 |
| 2;10 | 50% | 2 | 50% | 2 | 50% | 1 | 50% | 1 |
| 2;11 | 100% | 2 | 0% | 0 | 100% | 1 | 0% | 0 |
| Total | 47% | 47 | 53% | 53 | 33.3% | 36 | 66.7% | 72 |

In Eve's data, the formation of the sigmatic past tense with/without the presence of augmentation is mixed. The non-sigmatic past, on the other hand, is predominately

formed without an augment. In the first two recordings in particular, there are very few instances of past tense (1;7, 1;8), thus this is not very informative. Then, from the age of 1;9 to 2;0 the use of the sigmatic past with augment is higher than without augment (except in 1;11). From the age of 2;2 to 2;6 the sigmatic without augment is used more often than the sigmatic with augment. Finally, from the age of 2;9 to 2;11, there are only few instances of past tense use and little can be revealed from the results. In terms of the non-sigmatic past tense, the first three recordings provide very few instances of past tense verbal forms. From the age of 1;10 onwards, Eve's use of the non-sigmatic past without the augment is more prominent except at the ages of 1;11 and 2;5 where the use of augmented non-sigmatic past tense is in proportion with non-augmented ones (50%-50% and 55.5%-44.5% respectively). Lastly, the recordings at the ages of 2;9, 2;10 and 2;11 present very few non-sigmatic past tense verbal forms.

Eve and Maria form most of their verbs with an augment in the past tense. There is, however, a clear pattern seen in the tables above of the non-sigmatic past being formed more likely without an augment. In Eve's case, 66.7% of the non-sigmatic past has no augment morphology and with Maria, 80.4% of the verbs are formed without an augment.

In the sigmatic past, the proportion of presence/absence of the augment is better balanced. 47.9% of Eve's of sigmatic past verbs are formed with an augment and 37.2% of Maria's of sigmatic past tense verbs have an augment.

Augment, as mentioned earlier in this chapter, does not discriminate between sigmatic and non-sigmatic past tense; it applies to any verb regardless of whether or not its past tense is formed with the morpheme [-s] or not. Moreover, Mastropavlou (2006) suggests that when the augment rule²⁹ is applied, then this form is the most salient

²⁹ Stress shift and augment are considered to be a combined feature for the formation of past tense in MG (Mastropavlou, 2006).

version of the past tense. According to Mastropavlou, the [+/-augment] rule provides a perceptual salience factor for past tense formation. As a result, [+augment] presents higher salience than [-augment] past formation.

Based on this, the findings of this analysis suggest that sigmatic past tense is used both with and without augment, depending on the verb. The use of augmented and non-augmented sigmatic past forms varies in time, but overall both forms are used by Maria and Eve. The non-sigmatic past tense is mostly formed without augmentation. This result does not confirmed the prediction that [+augment] past tense forms are more salient and therefore easier to acquire, since both Maria and Eve used both augmented and non augmented past tense from the onset of the data collection.

A further analysis on regularity of past tense will reveal more on the issue of salience.

4.8.5 *Regularity*

As mentioned before, regularity (as opposed to augmentation) is part of a verb's individual configuration. As a result, a developmental analysis of past tense degree of regularity is not performed, i.e. the [+/- regular] will be analysed for all verbs together and not by age/recording. What is important to establish, however, is how many of the sigmatic and non-sigmatic past tense uses in Eve's and Maria's data are regular or irregular. The data will be analysed in terms of tokens for convenience and comparison purposes.

Following Mastropavlou's classification (Table 4.2) and her division of past tense into regular and irregular, the status of regularity³⁰ and salience will now be examined.

Table 4.10: Maria's [+/-regular] in sigmatic and non-sigmatic past tense

| Maria/Age | Sigmatic | | | | Non-sigmatic | | | |
|--------------|----------|----|-----------|----|--------------|---|-----------|----|
| | Regular | N | Irregular | N | Regular | N | Irregular | N |
| Total | 48.8% | 21 | 51.2% | 22 | 6.5% | 3 | 93.5% | 43 |

Table 4.11: Eve's [+/-regular] in sigmatic and non-sigmatic past tense

| Eve/Age | Sigmatic | | | | Non-sigmatic | | | |
|--------------|----------|----|-----------|----|--------------|---|-----------|-----|
| | Regular | N | Irregular | N | Regular | N | Irregular | N |
| Total | 63% | 63 | 37% | 37 | 5.5% | 6 | 94.5% | 102 |

Tables 4.10 and 4.11 present the distribution of regular/irregular sigmatic/non-sigmatic past tenses in Maria's and Eve's speech production.

The prediction is that [+irregular] past forms will have higher perceptual salience, if stress shift is involved as well as stem substitution because irregular forms are stored in the memory (Mastropavlou, 2006). From Tables 4.10 and 4.11, some of Mastropavlou's observed results are indeed confirmed, i.e. irregular verbs mostly form their past tense without the aspectual morphological marker [+s]. The pattern is not very apparent for the sigmatic past tense, since, especially in Maria's data, regular sigmatic

³⁰ The regularity of the past tense verbs was based on the *Tables of irregular verbs* from Holton et al. (1997; 169-175).

past is 48.8% and the irregular provide 51.2% of the examined production. The non-sigmatic past tense results confirm previous claims, since more than 90% of the cases involve an irregular verb for both Eve and Maria. Thus, the data appears to demonstrate that the lack of the aspectual morphological marker [-s] is closely tied to irregularity.

A chi-square statistical test was performed to examine the effect of regularity to the morphological distinction, i.e. sigmatic/non-sigmatic. The test reveals that the value in Pearson's chi-square is 77.289 in Eve's data and 20.206 in Maria's data. Both values are highly significant ($p < .001$), indicating that sigmatic morphology has had a significant effect on regularity. This significant effect means that sigmatic forms are mostly regular whereas non-sigmatic forms are mostly irregular, i.e. the pattern of regularity/irregularity in the sigmatic vs. non-sigmatic condition is significantly different.

To summarise some of the findings presented so far, the frequency of use of the sigmatic and non-sigmatic past tenses fluctuates over time. Sigmatic past tense forms appear to be used more often at an early age and the use of the non-sigmatic past tense increases at a later age. Augmentation and regularity appear to be involved more often with past tense forms with sigmatic marking, whereas non-sigmatic past forms often involve irregular verbs without augment.

Following Mastropavlou's classification of perceptual salience, the prediction is that irregular past tense without augmentation and without the aspectual marker -s- has the highest salience in the comprehension and production of past tense. This claim is not clearly supported from the analysis so far. The trace-back and frequency analyses presented in the following section will shed more light on this matter.

4.9 Trace-back procedure

Trace-back procedures are procedures designed to trace words or utterances (or strings of words) from a test file (usually the final recording) back to the rest of the data using matching operations as discussed in 2.4.1.1.

Trace-back methods were developed within usage-based models. Using corpora of speech production, trace-back methods identify multiword utterances and examine how these utterances are related to what the children has said before (Dabrowska and Lieven 2005; Lieven 2006; Lieven et al. 2003).

In the following section, I will discuss the *trace-back* procedure in more depth and how such an analysis offers additional insight for this study.

Usage-based theories suggest that evidence for productivity in trace-back procedures can be found in **novel** verbal forms. In this section I will examine *novelty* as a productivity criterion using a trace-back procedure (as applied previously to child and adult production data) for past tense.

The corpora used in this study were not originally designed for a trace-back analysis, so there are certain limitations on the way the data can be used for such a procedure. There is, however, scope on performing such analysis since there are no previous studies in MG examining usage-based theories and their predictions.

The trace-back procedure discussed in this section was designed following studies by Lieven, Behrnes, Spears and Tomasello, (2003) and Lieven et al., (2009) among others (see section 2.4.1.1).

4.9.1 Method

The procedure used in this study is much simpler than the trace-back system used in previous usage-based studies. It traces only words (past tense verbs) from a *later*

recording (if not the final one) back to the *previous recordings* of each child. The former will be called *test recording* and the latter *main corpus*. This terminology was adopted from Lieven, Behrnes, Spears and Tomasello (2003) and Lieven et al., 2009. The matching procedure in the present analysis is straightforward and there is no need for the use of deriving operations, as in the original study because matches are only identified for words and not phrases.

Through the trace back procedure, the following information can be established:

How many of the past verbal forms identified in the test file:

- a) Can be traced back to the child's previous files?
- b) Can be traced back to the adult's speech production?
- c) Are novel?

Past tense verbal forms were selected from the test recordings of each child. In Eve's data, the final recordings are very short and, there are disproportionate to the main corpus. In order to create more robust evidence for this procedure, recording number 10 (age 2;6) was used because it is larger in terms of speech production.

CLAN with COMBO commands were used for the identification of the matching verbal forms. These commands were also used to establish the exact phrase and disambiguate ambiguous cases, e.g. the imperative is homophonous to the 3rd singular past verbal forms:

djavase: 2ndSG Imperative/3rd SG Aorist (= read/he read)

For some verbs, two or more different verbal stems were checked to accommodate morpho-phonological variation e.g. the past tense of the verb *feug-o* (= I leave) is as follows:

e-fiġ-a, e-fiġ-jes, -efi-je, fiġ-ame, fiġ-ate, e-fiġ-an
 (= I left, you left, he/she/it left ... etc.)

where the following stems were analysed: *e-fiġ-, e-fiġ-, 0-fiġ-*.

Past tense verbs from the test file were checked for matches in all previous recordings, firstly in the child's tiers and then in the adults speech. Where more than one adult was involved in the recording, all participants were checked using a COMBO command with the coding name, e.g.:

*combo +t*AD1 +s*leroġhik* eve6.cha*

*combo +t*AD2 +s*leroġhik* eve6.cha*

*combo +t*AD3 +s*leroġhik* eve6.cha*

The procedure also followed the original classification between sigmatic and non-sigmatic use of past tense, as originally proposed in this analysis.

Table 4.12: Ratio of analysed utterance between test recording and main corpus

| | Main Corpus | Test (recording) |
|--|-------------|------------------|
| Maria - No of analysed utterances | 1261 | 160 (12.68%) |
| Eve - No of analysed utterances | 6916 | 379 (5.48%) |

Table 4.12, shows the ratio of analysed utterances between the test recording and the main corpus (based on Lieven et al., 2009).

Matches in the trace-back analysis were performed in two sets:

Set 1: *Exact* match

Set 2: *Closest* match (in brackets)

Exact match is when the matched past tense is in the same person as the target past, whereas *closest* match refers to other persons of the same past tense verbs. *Closest* matches are given in brackets next to the exact match. Past tense which is not matched with any previous occurrence (neither in the child's nor in the adults' speech) is considered to be *novel*.

The prediction is that children's word production can be traced back in their speech and this is also often the case with the adults. An obvious limitation of this trace back procedure is that the analysed utterances represent only a very small percentage of what the child listens to and produces on a daily basis.

Tables 4.13 and 4.14 below present all past tense matches identified in the children's and adults' speech. The tables were ordered by number of matches in the main corpus.

Table 4.13: Maria's matched past tense in the main corpus and in adults' speech

| Target word in test (sigmatic) | Number of exact matches in main corpus (closest match) | Number of matches in adult speech |
|---|---|--|
| Djavasa (= read/1sg) | 3 | (1) |
| Epjase (= held/3sg) | 3 | 0 |
| Eklise (= closed/3sg) | 2 (1) | (1) |
| Teliose (= finished/3sg) | 2 | 0 |
| Echise (= poured/3sg) | 2 | 0 |
| Teliosa (= finished/1sg) | (2) | 0 |
| Eklisa (= closed/1sg) | 1 (2) | 1 |
| Eftiakse (= made/3sg) | 0 | (1) |
| Kukulosa (= covered/1sg) | 0 | 0 |
| Target word in test (non-sigmatic) | Number of exact matches in main corpus (closest match) | Number of matches in adult speech |
| Ipa (= said/1sg) | 11 (1) | 1 (3) |
| Pije (= went/3sg) | 2 (1) | 0 |
| Pire (= took/3sg) | 1 (2) | 0 |
| Pigane (= went/3PL) | 1 (2) | 0 |
| Evgala (= removed/1sg) | 1 | 0 |
| Anevike (= got up/3sg) | 0 | 0 |

Table 4.14: Eve's matched past tense in the main corpus and in the adults' speech

| Target word in test (sigmatic) | Number of exact matches in main corpus (closest match) | Number of matches in adult speech |
|---|---|--|
| Epese (= fell/3sg) | 14 (9) | 27 (4) |
| Epesa (= fell/1sg) | 7 (16) | (31) |
| Chalase (= got ruined/3sg) | 2 (5) | 14 (5) |
| Pesame (= fell/1PL) | (24) | (31) |
| Djavasa (= read/1sg) | 2 | 1 (3) |
| Majirepse (cooked/3sg) | 0 | (2) |
| Ekripsa (= hid/1sg) | 0 | (1) |
| Egrapse (= wrote/3sg) | 0 | 0 |
| Filise (= kissed/3sg) | 0 | 0 |
| Travikse (= pulled/3sg) | 0 | 0 |
| Klotsise (= kicked/3sg) | 0 | 0 |
| Target word in test (non-sigmatic) | Number of exact matches in main corpus (closest match) | Number of matches in adult speech |
| Irthe (= came/3sg) | 6 | 12 (1) |
| Efije (= left/3sg) | 7 (1) | 8 |
| Ipe (= said/3sg) | 5 (4) | 15 (62) |
| Evale (= put/3sg) | 3 (2) | 6 (15) |
| Pira (= took/1sg) | 3 (2) | 5 (23) |
| Ipa (= said/1sg) | 3 (6) | (77) |
| Piga (= went/1sg) | 2 (8) | 1 (40) |
| Pigame (= went/1PL) | 1 (8) | 24 (17) |

| | | |
|--------------------------------|-----|---------|
| Ekatsa (= sat/1sg) | 1 | (5) |
| Ipame (= said/1PL) | (9) | 17 (60) |
| Ide (= saw/3sg) | (3) | 2 (18) |
| Iche (= had/3sg) | 0 | 31 (16) |
| Bike (= came in/3sg) | 0 | 7 (4) |
| Chathike (= got lost/3sg) | 0 | 0 |
| Lerothika (= got soiled/1sg) | 0 | 0 |
| Padreftike (= got married/3sg) | 0 | 0 |
| Dithika (= got dressed/1sg) | 0 | 0 |
| Dithike (= got dressed/3sg) | 0 | 0 |
| Edire (= hit/3sg) | 0 | 0 |
| Kriftika (= was hidden/1sg) | 0 | 0 |

In Tables 4.13 and 4.14 the matching information is presented. The matches should exhibit the predicted pattern, i.e. that most words used at the test recording can be traced back into the child's or the adult's speech.

The tables show the matches (exact and closest) of verbs in past tense (sigmatic and non-sigmatic) traced-back from the test recording to the children's previous speech production. Novel past tense, if not traced back in the child's speech often has no matches in the adults' speech e.g.

filise (= kissed/3sg)

klotsise (= kicked/3sg)

dithika (= got dressed/1sg)

kukulosa (= covered/1sg)

anevike (= got up/3sg)

This is observed more in Eve's data than in Maria's. Adult participants were not very interactive with Maria during her recording. This contextual information can explain the very low matching occurrences of Maria's target past tense in the adults' speech.

Often, words are traced-back successfully in the adults' speech but not in the children's main corpora:

ekripsa (= hid/1sg)

majirepse (= cooked/3sg)

iche (= had/3sg)

eftiakse (= made/3sg)

These matches are assumed to be newly used by the children and hence just recently acquired.

Successfully traced-back words, according to usage-based theories, reveal that children build novel construction on what they know and use already. The claim is also that the more a child uses a construction or a verb, the more they use it and know how to use it correctly. Thus, frequent words increase in frequency after being used more. Children's knowledge and use is based on what they hear in their immediate environment. Productive use of what children learn only comes later after the age of 3.

It would be useful at this point to propose a new criterion of productivity based on the presence of *closest* matches. Closest matches, as mentioned above, involve the same verb, used in a different person. That is, *pesame/1pl* (= we fell) is first used in the

test recording, but there are 24 closest matches of the same verb in different persons in the child's previous production and 31 occurrences in the adults' speech. This can be an indicator of productivity of the verb's past tense paradigm since the child is able to produce a novel verb form based only on information from other verbs or general knowledge of verb conjugation. Thus, morphological information from other verbs may be used to produce novel forms of a verb, but this does not necessarily happen on a verb-by-verb basis.

Usage-based theories support the idea that children near the end of their 2nd year start showing some signs of paradigmatic categorisation of verbs. In support of the usage-based theory, there is evidence from the analysis that children produce previously used verbs because they are familiar with the verbs and their combinatorial characteristics.

Examples of the higher frequency verbs are:

- *chalase* (= got ruined/3sg) is used twice in previous recordings (exact match) and 5 times in other persons (closest match)
- *epesa* (= fell/1sg) is used 7 times in previous production and 16 times in other persons
- *irthe* (= came/3sg) is used 6 times previously (no closest matches)
- *efije* (= left/3sg) is used 7 times before and once in another person
- *eklise* (= closed/3sg) is used 2 times before and once in another person
- *ipa* (= said/1sg) is used 11 times in the main corpus and once in another person.

That is, children use the verbs they know and often new uses of a given verb replicate previous uses with only one addition or a small modification. Usage-based accounts

would suggest that this is done on an item-specific basis. However, there is limited evidence here to support this claim because of the limitations of the trace-back procedure used. Trace-back procedures used in usage-based studies (Lieven et al., 2009; Dabrowska and Lieven 2005; Lieven 2006; Lieven et al. 2003) are designed to identify two types of component units in the preceding corpus i.e. fully lexically-specified string of one or more words and schemas with slots. Matching operations are consequently used to identify matching utterances in the main corpus using distributional and semantic criteria. In this study, however, it was not possible to employ such semantic matching criteria in the trace-back procedure because the data collection was not originally designed for this purpose.

In summary, the results of the data analysis provide evidence against item-by-item learning, and thus, against usage-based models. The prediction of usage-based theories was that children's past tense can be traced back in their earlier speech and also in the adults' speech. The findings of the trace-back procedure performed in this section suggest that children often use verbs that they used before building on their knowledge of these verbs. Such use can also be traced-back in the adult speech suggesting that children's speech is closely related to the input they are exposed to. The presence of novel verbs, however, provides evidence that children use new verbs based on previous knowledge they have; such knowledge is transferred by analogy. The assumption of usage-based models that children use only what they used before and what they know well from their previous production is not borne out from this analysis and therefore this provides counterevidence for the usage-based models.

4.9.2 Frequency

Following the above findings, further analyses of frequencies of the verbs used in the corpora were made to indicate if some of these verbs are used more frequently than others in the adult language. For this purpose I used the resources of the Hellenic National Corpus available from the Institute for Language and Speech Processing³¹.

The Hellenic National Corpus contains more than 47.000.000 words of written texts from books, mass media, and the World Wide Web. The corpus provides statistical data for whole sentences, parts of speech, words and lemmas. It is assumed that the statistical information provided also corresponds to oral language to some extent (Mastropavlou, 2006).

The statistical information of the Hellenic National Corpus offers two types of search, namely, *word* or *lemma*. I will perform the analysis using the *word* statistical search rather than the *lemma* one for the following reason: *lemma* search provides information about the verb in 1st singular present tense, whereas *word* search provides information for the specific verbal form used in the children's data, i.e. the past tense.

The sigmatic and non-sigmatic past tense from Eve's and Maria's corpora from Tables 4.13 and 4.14 were divided in high and low frequency of use in the following Tables 4.15 and 4.16. The division was based solely on matching information to the child's production but not to the adults'. The adults' matching information will be taken into consideration in the discussion of frequency.

³¹ <http://www.ilsp.gr/index.php>

Table 4.15: **High** frequency used past tense in Maria's and Eve's speech (divided into sigmatic and non-sigmatic)

| Sigmatic - Maria | Used | Non-sigmatic - Maria | Used |
|-----------------------------|-------------|-----------------------------|-------------|
| Epjase (= held/3sg) | 3 | Ipa (= said/1sg) | 11 |
| Djavasa (= read/1sg) | 3 | Pije (= went/3sg) | 2 |
| Eklise (= closed/3sg) | 2 | Pire (= took/3sg) | 1 |
| Teliose (= finished/3sg) | 2 | Pigane (= went/3pl) | 1 |
| Echise (= poured/3sg) | 2 | Evgala (= removed/1sg) | 1 |
| Eklisa (= closed/1sg) | 1 | | |
| | | | |
| Sigmatic - Eve | Used | Non-sigmatic - Eve | Used |
| Epese (= fell/3sg) | 14 | Efije (= left/3sg) | 7 |
| Epesa (= fell/1sg) | 7 | Ipe (= said/3sg) | 5 |
| Chalase (= got spoiled/3sg) | 2 | Irthe (= came/3sg) | 6 |
| Djavasa (= read/1sg) | 2 | Evale (= put/3sg) | 3 |
| | | Pira (= took/1sg) | 3 |
| | | Ipa (= said/1sg) | 3 |
| | | Piga (= went/1sg) | 2 |
| | | Pigame (= went/1pl) | 1 |
| | | Ekatsa (= sat/1sg) | 1 |

Table 4.16: **Low** frequency use of past tense in Maria’s and Eve’s speech (divided in sigmatic and non-sigmatic)

| Sigmatic – Maria | Used | Non-sigmatic – Maria | Used |
|--------------------------|-------------|--------------------------------|-------------|
| Kukulosa (= covered/1sg) | 0 | Anevike (got up/3sg) | 0 |
| Eftiakse (= made/3sg) | 0 | | |
| Teliosa(= finished/1sg) | 0 | | |
| | | | |
| Sigmatic – Eve | Used | Non-sigmatic – Eve | Used |
| Egrapse (= wrote/3sg) | 0 | Lerothika (= got soiled/1sg) | 0 |
| Travikse (= pulled/3sg) | 0 | Dithika (= got dressed/1sg) | 0 |
| Pesame (= fell/1pl) | 0 | Ipame (= said/1pl) | 0 |
| Majirepse (= cooked/3sg) | 0 | Ide (= saw/3sg) | 0 |
| Ekripsa (= hid/1sg) | 0 | Iche (= had/3sg) | 0 |
| Filise (= kissed/3sg) | 0 | Bike (=came in/3sg) | 0 |
| Klotsise (= kicked/3sg) | 0 | Chathike (got lost/3sg) | 0 |
| | | Padreftike (= got married/3sg) | 0 |
| | | Dithike (= got dressed/3sg) | 0 |
| | | Edire (= hit/3sg) | 0 |
| | | Kriftika (= was hidden/1sg) | 0 |

Tables 4.15 and 4.16 show verbs from Maria’s and Eve’s corpora with high/ low usage frequency in the children’s speech production and divided into sigmatic and non-sigmatic.

The Hellenic National Corpus provides the following statistical information for each of the verbs used by the children, presented in the table above:

Table 4.17: Frequencies of all past tense verbs from the Hellenic National Corpus in descending order (sigmatic in italics)

| Word | Occurrences | Frequency (‰) | l(ow) ³² h(igh) |
|---------------------------------|--------------------|----------------------|-------------------------------|
| Iche (= had/3sg) | 71672 | 1.5245 | l |
| Ipe (= said/3sg) | 35201 | 0.7487 | h |
| Pire (= took/3sg) | 7839 | 0.1667 | h |
| Ipa (=said/1sg) | 5709 | 0.1214 | h |
| Irthe (= came/3sg) | 5243 | 0.1115 | h |
| Pije (= went/3sg) | 4021 | 0.0855 | h |
| Ide (= saw/3sg) | 3345 | 0.0711 | l |
| Evale (= put/3sg) | 2508 | 0.0533 | h |
| Bike (= came in/3sg) | 2508 | 0.0533 | l |
| Efije (= left/3sg) | 2363 | 0.0503 | h |
| <i>Egrapse</i> (= wrote/3sg) | 2327 | 0.0495 | l |
| <i>Epese</i> (= fell/3sg) | 2305 | 0.049 | h |
| <i>Eklise</i> (= closed/3sg) | 2123 | 0.0452 | h |
| Ipame (= said/1pl) | 1917 | 0.0408 | l |
| <i>Teliose</i> (= finished/3sg) | 1241 | 0.0264 | h |
| <i>Epjase</i> (= held/3sg) | 1006 | 0.0214 | h |
| Piga (= went/1sg) | 925 | 0.0197 | h |

³² This column presents high and low frequency items as produced in Maria's and Eve's data.

| | | | |
|---------------------------------|-----|--------|---|
| Pira (= took/1sg) | 893 | 0.019 | h |
| Anevike (= got up/3sg) | 865 | 0.0184 | l |
| Chathike (= got lost/3sg) | 842 | 0.0179 | l |
| Travikse (= pulled/3sg) | 568 | 0.0121 | l |
| Pigame (= went/1pl) | 518 | 0.011 | h |
| Median Threshold | | | |
| Djavasa (= read/1sg) | 509 | 0.0108 | h |
| Eftiakse (= made/3sg) | 471 | 0.0100 | l |
| Padrefitike (= got married/3sg) | 322 | 0.0068 | l |
| Chalase (= got spoiled/3sg) | 243 | 0.0052 | h |
| Evgala (= removed/1sg) | 163 | 0.0035 | h |
| Filise (= kissed/3sg) | 152 | 0.0032 | l |
| Eklisa (= closed/1sg) | 123 | 0.0026 | h |
| Epesa (= fell/1sg) | 113 | 0.0024 | h |
| Teliosa (= finished/1sg) | 112 | 0.0024 | l |
| Pigane (= went/3pl) | 73 | 0.0016 | h |
| Pesame (= fell/1pl) | 73 | 0.0016 | l |
| Dithike (= got dressed/3sg) | 67 | 0.0014 | l |
| Klotsise (= kicked/3sg) | 57 | 0.0012 | l |
| Edire (= hit/3sg) | 28 | 0.0006 | l |
| Echise (= poured/3sg) | 25 | 0.0005 | h |
| Ekatsa (= sat/1sg) | 21 | 0.0004 | h |
| Majirepse (= cooked/3sg) | 16 | 0.0003 | l |
| Ekripsa (= hid/1sg) | 12 | 0.0003 | l |

| | | | |
|---------------------------------|----|--------|---|
| Kriftika (= was hidden/1sg) | 10 | 0.0002 | 1 |
| Dithika (= got dressed/1sg) | 5 | 0.0001 | 1 |
| Lerothika (= got soiled/1sg) | 0 | 0.0000 | 1 |
| <i>Kukulosa</i> (= covered/1sg) | 0 | 0.0000 | 1 |

Table 4.17 shows the most and least used (marked with l(ow) and h(igh)) of the past tense verbs from Eve's and Maria's corpora. For convenience the sigmatic past is given in italics.

For the trace-back procedure, it may be useful to identify a frequency threshold between high and low frequency past tense. The median of the occurrences of the high and low frequencies of the verbs in Table 4.17 was calculated and it is 513.5 (0.0109%). Based on this median I propose the threshold between high and low frequency past tense to be around 0.0109% (<513 occurrences). From the Table 4.17 the division between high and low frequency items is not very clear in relation to the children's usage. In fact, 8 out of 22 children's high frequency verbs have low frequency in the adult language and 8 out of 22 children's low frequency verbs have high frequency in the adult language according to the proposed threshold. Therefore the majority of high frequency past tense verbs appear to be frequently used in children's speech and the majority of low frequency past tense appears to be less frequently used in their early speech. There are however, instances of low frequency - in the adult language - verbs, used frequently in the children's speech, such as *iche* (= *had/3sg*), *ide* (= *saw/3sg*), *bike* (= *came in/3sg*), *egrapse* (= *wrote/3sg*).

Additionally, based on the information presented in Tables 4.13 and 4.14, the matches to the adults' speech also reveal an interesting pattern, i.e. most of the low frequency past tense have no matches with the adults' data, whereas most of the high

frequency items do have matches, i.e. *iche* (= *had/3sg*), *epese* (= *fell/3sg*), *irthe* (= *came/3sg*), *pigame* (= *went/1pl*).

To summarise, the high frequency past tense forms in adults' speech are used more often in children's speech and it can also be traced-back to the children's previous production and to the adults' speech as shown in Tables 4.13 and 4.14. Low frequency past tense forms, on the other hand, are scarcely used and cannot be traced-back to either the children's previous speech or to the adults' speech. Children appear to use high frequency verbs more often than low frequency verbs. This may be because they hear these items more often and they are more likely to use the verbs they are familiar with. The data supports the notion of frequency and frequent use of recurrent items as having a significant effect in children's speech, but does not provide any support for the item-specific claim of usage-based accounts.

4.10 Summary and discussion

The aim of this chapter was to investigate the acquisition of past tense in Modern Greek by analysing Maria's and Eve's speech. The analysis intended to answer two research questions. Firstly, from a nativist point of view, whether there is a difference between sigmatic and non-sigmatic past tense in Maria's and Eve's speech. Secondly, from a usage-based point of view, whether the children's past tense can be traced back to their earlier speech production and to the adult language and what effect frequency has on children's use of past tense.

The data analysis presented here has brought to light certain patterns of past tense use which are visible only by focusing on very early stages of language development. It has been suggested that children perform better in the comprehension and production of sigmatic past tense in existing and novel verbs (Delidaki and

Varlokosta, 2003; Stathopoulou, 2009; Stavrakaki and Clahsen, 2009) because sigmatic past tense is rule-based in MG. The findings of the current study suggest that the use of sigmatic and non-sigmatic past tense varies over time; sigmatic past tense is used more often at an early age while the use of non-sigmatic past tense increases soon after. Finally sigmatic past increases again. There is a U-shaped fashion on the development of sigmatic past tense. This is also confirmed through the MLUw analysis for Eve's speech production. In terms of single and dual mechanism, these findings provide support for the dual mechanism of acquisition of past tense, where the sigmatic past is the default past tense and hence acquired on a rule-based manner while irregular past tense is learned on a one-to-one basis and stored in the lexicon.

According to Mastropavlou (2006), the degree of salience of past tense determines children's degree of learnability, i.e. the higher the salience of the past tense the better children will perform on comprehending and producing this form. In order to account for Mastropavlou's proposal the augment and regularity were examined in the production of past tense in the children's corpora. The findings of such analysis suggest the augmentation and regularity involve mostly sigmatic past tense while non-sigmatic past tense is more often irregular and without augment. The prediction that irregular past tense is more salient (Mastropavlou, *ibid.*) was not borne out from the analyses.

Finally, in an attempt to account for usage-based approaches to language acquisition, suggesting that language structure emerges from language use, a trace-back procedure was designed and used to examine how children's past tenses at their last recording are related to what the children has said before. Also frequency of those items in adult language was analysed in order to establish the relation between frequency and use. The findings have revealed that high frequency past tense forms in adults' language are used more often in children's speech and it can be traced-back to the children's

previous production and to the adults' participants' speech. In contrast, low frequency past tense forms are scarcely used and cannot be traced-back to either the children's previous speech or to the adults' speech. These results support, to some extent, the suggestion that children use more the words and utterances that they already know and their new production is based on those early structures. Finally, the analysis of productivity of the use of novel verbs in past tense (verbs that cannot be traced-back) provides evidence against the prediction that children learn verb on an item-by-item basis.

5 ROOT INFINITIVES

5.1 Aims of the chapter

This chapter aims to investigate the use of finite and non-finite verbal forms in early Greek and elucidate the status of RIs based on the analyses proposed in previous research through the analysis of new data.

There are two sets of questions to be addressed. The first set of questions is related to the categories of finiteness and inflection:

Do children use non-finite (bare) verbal forms in main contexts in their early speech production; if yes is this evidence for two stages in the acquisition of early Greek?

The second set of questions originates from a usage-based approach and investigates the role of the input frequency in language acquisition as well as the effect of the morphological complexity of MG:

Are children's early non-adult constructions (RIs) the result of frequency of some structures from the adult input or simply the result of performance inaccuracy (reduced production)?

Is the RIs stage the result of morphological ambiguity of the MG verbal system?

The first set of questions will explore the production of finite and non-finite clauses in Maria's and Eve's corpora, building on previous analyses for RIs discussed in

the first part of this chapter. These analyses will investigate the status of the modal particles, aspect and telicity, 3sg-suffix, and past tense. It will also consider evidence for the presence of developmental stages and in what stage Maria's and Eve's speech development can be positioned. Based on the findings of these analyses this section will test out the status of the RI stage in the child Greek. Based mainly on VVR's study, the predictions for the first question can be summarised as:

- Maria's and Eve's speech develops in two stages. During the first stage, *-i* forms are overused. Such forms are more prominent with the perfective than with the imperfective stem. *-i* forms also occur with incorrect agreement i.e. in non-3sg contexts and usually with null subjects contrary to finite verbs, where both null and overt are observed.
- During stage II children's use of non-finite non-adult like forms reduces and soon such forms disappear.

The predictions to the second set of questions focus on the following topics:

- Frequency and adult input: frequently used items in the input will be frequently used in the children's speech production.
- Morphological ambiguity of the language learned: the degree of the RI stage in children's production is the result of the degree of ambiguous non-finite verbal forms in a given language, i.e. the more ambiguous is the verbal system the higher the rate of RIs will be.

5.2 Introduction

A particular area of grammar where children seem not to converge immediately on to the adult target concerns the requirement that matrix clauses are finite. Children acquiring various languages, including German, Dutch, French, and Swedish show a robust RI³³ effect. Children use infinitives in main clauses (root context) which would be ungrammatical in the adult language as shown below:

| | | |
|----------------|---|----------------------------|
| German | du das haben <i>you that have-INF</i> | (Poeppel and Wexler, 1993) |
| Dutch | Op kist zitten <i>on box sit-INF</i> | (Salustri and Hyams, 2003) |
| French | Michel dormir <i>Michel sleep-INF</i> | (Pierce, 1992) |
| Swedish | Gubbe vara dar <i>old-man be-INF there</i> | (Platzack, 1992) |

The age at which the phenomenon occurs is roughly between **2;0** and **2;6**. During this time, RIs constitute the 30% - 70% of the children's verbal utterances. The remaining verbal utterances are adult-like finite clauses. Thus, RIs occur side-by-side with well-formed finite verbs. RIs typically have a modal interpretation, their reference is not free (Hoekstra and Hyams, 1998).

³³ *Optional Infinitives*, (term due to Wexler, 1994) is also used to designate the same phenomenon.

The RIs phenomenon is not a universal property of child language. Children acquiring Italian, Spanish, Basque and Catalan do not show a RI stage; RIs do not occur in languages with syntactically strong agreement. Even though children acquiring rich agreement languages (null subject languages) do not exhibit a RI stage, they do produce bare verbal forms (Hyams, 2002). Such bare forms are usually infinitive in main clauses or in the case of Italian participles unsupported by an auxiliary:

| | | |
|----------------|---|------------------------|
| Italian | Babbo vedere la moto <i>daddy see-INF³⁴ the motorbike</i> | (Guasti 1994) |
| Spanish | Este tapar <i>this one cover-INF</i> | (Perales et al., 2004) |
| Basque | Hartu Ana! <i>take-INF Ana</i> | (Liceras et al., 2006) |
| Catalan | Sortir <i>come out-INF</i> | (Perales et al., 2006) |

English-speaking children also go through a stage, analogous to the RIs stage, in which a high percentage of bare forms, i.e., uninflected verbs is found. However, it has been suggested (Hyams, 2002) that the early English bare verb structure does not typically have a modal interpretation like the other RIs/RIs equivalents do. The English

³⁴ Some studies (Varlokosta, 2002; Hyams, 2002) suggest that children learning Italian produce both a bare past participle lacking the auxiliary (Presa Checco campana = *taken-PRTC Checco bell*) and an infinitive without a main verb. The suggestion is that RIs and bare participles may be manifestations of a more general root non-finite stage.

bare verb has either a present tense meaning or less frequently, a past tense meaning (Hyams, 2002).

English Eve sit floor
 Papa want apple

(Hyams, 2002)

There are two important properties regarding the RI stage. Firstly, during this stage children produce both finite and non-finite verbs in root clauses. In fact, Wexler (1994) argues that children know the difference between finite and non-finite clauses, i.e. in terms of movement, children know that finite verbs raise and non-finite verbs do not. A second important property related to the RI stage is that when finite forms are used they are used correctly, i.e. despite the rather limited agreement paradigm there are no agreement errors (Poeppel and Wexler, 1993). Both of these facts indicate knowledge of the adult grammar despite the suggested unproductive use of agreement at this stage. The idea of early knowledge of inflectional morphology is formalised by Wexler (1998) as the hypothesis of *Very Early Knowledge of Inflection (VEKI)* and by Hoekstra and Hyams (1995) as the hypothesis of *Early Morphosyntactic Convergence (EMC)*.

One of the most important findings emerging from the investigation of RIs in early grammar is that the subjects of these sentences are typically null, while the subjects of finite clauses occurring during the same period are typically overt. The high rate of null subjects in RIs contexts supports the claim that RIs are indeed non-finite. As non-finite clauses, they provide a licensing context for the null subject, analogous to the

situation in the respective adult languages, which license PRO in embedded infinitival clauses:

Andrew tried PRO to run away

Moreover, various studies have dealt with the semantic properties and meaning of RIs across the different languages and their similarities. It has regularly been observed that there is a constraint on the aspectual nature of RIs in languages such as Dutch and French, which is that only eventive verbs appear as RIs, while stative verbs typically require finiteness. In Hoekstra and Hyams (1998) this is referred to as the “eventivity constraint” on RIs. A second finding is that RIs typically receive a modal interpretation.

Although MG is a null subject language and has rich agreement, it has been suggested that children acquiring MG also go through an OI (Optional Infinitive) stage (Hoekstra and Hyams, 1995, 1998; Varlokosta, Vainikka and Rohrbacher, 1996, 1998; Hyams, 2001, 2002; Varlokosta, 2002), but the dataset used for this claim is very limited. Therefore, it is unclear whether or not such stage exists and for how long.

The current chapter examines the status of the RIs equivalent in Modern Greek and assesses the proposed evidence for RIs through a series of analyses of one new corpus of two children.

5.3 Analyses on RIs with Continuity and Maturation accounts

There are two main accounts of language acquisition within the generative framework, i.e. the Continuity and the Maturation accounts discussed in what follows. Under the Continuity account, there are two views: the Strong Continuity or Full Competence

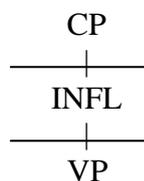
Hypothesis (FCH) suggesting that functional categories are present and fully specified from the very beginning in children's speech, and the Weak Continuity claiming that functional categories are present but underspecified. Under the Maturation account there are also two views, i.e. the Maturation or Structure Building Hypothesis (SBH) claiming that functional categories are not available in the early grammar but mature over time, and the Prefunctional Grammar model proposing a prefunctional grammar in which the functional module is an independent entity in the language faculty, subject to maturation in children's early speech.

According to the Strong Continuity or Full Competence Hypothesis account, all functional categories are present from the beginning. There are two flavours of this hypothesis. The first suggested that functional projections are not only present from the beginning but also fully specified (Boser et al., 1992; Poeppel and Wexler, 1993). Despite the absence of functional material, the presence of syntactic movement in early language is taken as an argument for the presence and full specification of functional categories.

A second variant of the FCH, the Weak Continuity view, claimed that functional categories are present but underspecified for their features (Hyams, 1992; Hoekstra and Hyams, 1995; Schutze and Wexler, 1996). The underspecification approach explained thus the absence of the full range of a paradigm in the early language. Within this view, RIs are the result of the underspecification of some functional category. Rizzi's (1993, 1994) view is also a version of the underspecification hypothesis. Rizzi (ibid.) proposed that children's phrases are reduced phrases in which a truncation mechanism has been applied.

According to Rizzi, all adult phrases are CPs. Children's phrases, however, can violate this axiom, applying this truncation mechanism, which cuts some categories

below the CP layer. Children use a more ample variety of root categories, the more simple ones, namely NP, PP, AdjP (non finite) VP.



Rizzi (*ibid.*) suggested that truncation is responsible for the presence of RIs in children's language. Rizzi's truncation theory has been abandoned in favour of more advanced propositions for the status of RIs, as we will see below.

According to the second account, the Maturation or Structure Building Hypothesis, functional categories are not available in the early grammar but mature according to a programme prescribed by UG (Radford, 1990; Tsimpli, 1996). There are weaker versions of this hypothesis proposed by Clahsen (1991), and Clahsen and Penke (1992).

The absence of morpholexical material associated with functional projections is taken as a strong argument for the Maturation or SBH. However, it is not the case that morpholexical material is totally absent from early language. Under the Maturation or SBH, RIs are taken as evidence for the lack of a functional category, namely the Inflection Phrase.

Another hue of the Maturation hypothesis is discussed in Tsimpli (1992/1996) namely the prefunctional grammar model. Tsimpli (*ibid.*) claimed that there is a prefunctional stage in children's language suggesting the absence of functional projections in the child's grammar.

Tsimpli assumed that functional categories determine linguistic variation in terms of parameterisation. Within the grammatical model presented, the total of

functional categories constitutes an independent module, independent in the language faculty. This is the Functional Module (FM), distinct from the lexical module, which includes lexical categories, like noun, verb, and adjective.

According to Tsimpili, UG principles are always available, but the functional module is subject to maturation and hence not available initially (18-24 months of age). The fact that the FM is subject to maturation explains why initially children omit functional material, like articles, auxiliaries, etc. This stage is referred to as the prefunctional stage, in which children do not use functional categories.

The notion of Maturation within language acquisition theory has been supported by various researchers (Felix, 1984; Borer and Wexler, 1987/1988; Guilfoyle and Noonan, 1992; Radford, 1988/1990). The hypothesis that there is a maturation process which affects the language development is plausible. The fact that certain properties of a biologically determined programme have to be available in specified timing conforms to the general idea that biologically determined development processes are restricted by maturation. In terms of phrase structure during the prefunctional stage, the Maturation Hypothesis concluded that functional categories consist only of projections of the substantive (lexical) categories.

Moreover, according to Tsimpili, children's verbal phrases include only a VP. The subject is generated within the VP projection since the IP projection does not exist yet. Consequently, the elements associated to the IP projection, in particular modals, and auxiliaries, are absent from the children's speech production. The absence of functional categories in the phrasal structures has a number of consequences regarding the linguistic availability of null subjects, the absence of movement as substitution processes, the absence of case assignment, and the possibilities associated with the word order properties.

What Tsimpli considered to be crucial evidence for her proposal of a prefunctional stage was the presence of cases where the agreement morpheme on the verb does not agree with the subject. This showed, according to Tsimpli, that the two elements are not in a proper configurational relation of Spec-Head agreement. Tsimpli claimed that what appeared to be agreement errors observed in child Greek are not in fact agreement errors but instances of non-agreement. Children do not make agreement errors; they either use agreement correctly or avoid it altogether.

To summarise, the principal assumptions of Tsimpli's theory, are the availability of the UG principles via the language acquisition processes and the non-availability of functional categories, due to maturation relations.

The theory of the prefunctional stage is important to this study since it has been proposed in various studies, discussed below, that the deficiency of agreement, and null-subjects characteristic of the RIs, have as immediate consequence the lack of functional categories in the children's speech. It is claimed that RIs can be explained by the lack of functional categories from children's early grammar, which consequently undergoes maturation in order to be adjusted according to the adult model.

5.4 Usage-based account of RIs

RIs are often used as an argument against usage-based acquisition models that consider input and frequency to be an important factor of learning. The rationale behind this assumption is that RIs being ungrammatical constructions are not present in the input since they are not produced by adults.

Usage-based approaches, however, account for the phenomenon of RIs in children's speech, as it will be shown in the following sections.

5.4.1 Simulation of RIs using MOSAIC

The main volume of work around RIs from a usage-based point of view was conducted by Freudenthal, Pine and Gobet (2004, 2005, 2006, 2009) using the model of syntax acquisition in children (MOSAIC).

MOSAIC is a simple discrimination net (an n-ary tree) which is headed by a root node and where more nodes are added as the model is shown more utterances (Freudenthal et al., 2004, 2005, 2006). MOSAIC learns from orthographically coded input, with whole words being the unit of analysis while utterance production involves outputting all the utterances the model has encoded (Freudenthal et al., *ibid.*).

MOCAIC incrementally learns and stores utterances that are inputted to it. MOSAIC also produces novel utterances using a mechanism which links words or phrases occurring in similar context; thus, two linked elements can be substituted for each other (Freudenthal et al., *ibid.*).

Freudenthal et al. (*ibid.*) through MOSAIC assume that language learning is a performance limited process which is predominantly influenced by the most recent elements uttered.

MOSAIC has been used to simulate children's RIs by omitting sentence-internal elements (Freudenthal, et al., 2005). The model learns from both the left and right edge of the utterance, and associates utterance-initial and utterance-final phrases (Freudenthal, et al., 2005). RIs are suggested to be a natural result of primacy and recency effects in children's learning mechanism which eventually disappear as the length of utterance increases. This, in essence, is the result of frequency effects, i.e. the input drives the learning mechanism.

5.4.2 *Kupisch and Rinke (2007)*

Kupisch and Rinke's study (2007) suggested that RIs is a temporary phenomenon in children's language that can be explained as the learning of the inflectional properties of a given language whose verbal paradigm has certain ambiguities.

The study analysed individual transcripts of children's speech production in Italian, Brazilian Portuguese, French, German and English, aged between 2;0 and 2;2 (Kupisch and Rinke, *ibid.*).

Children's data were classified based on the following features:

- unambiguously finite verbs
- non-finite target-like verbs (licensed by preceding auxiliaries, modals or prepositions)
- non-finite non-target-like verbs (gerunds, participles and infinitives which are "bare", i.e. not preceded by auxiliaries, modals or prepositions)
- morphologically ambiguous forms (finite verbs and categories other than verbs homophonous with infinitives)
- copula omissions.

(Kupisch and Rinke, 2007:91)

The tense morphological systems of each language were also examined and the overall findings suggest that children learning languages, such as Italian and Brazilian Portuguese, whose tense morphology is the most explicit, produced the lowest number of non-target-like infinitives, i.e. RIs (Kupisch and Rinke, *ibid.*). English and German learning children, on the other hand, produced more RIs being the results of high amount of ambiguities between finite and non-finite verbs in their respective languages.

Kupisch and Rinke (ibid.) maintained that there is a relation between the quantity of RIs in children's language and the morphological properties of the language they are exposed to. Therefore, in languages in which tense can be expressed with a variety of verbal forms, the children's rate of RIs is very low and the acquisition process is faster, whereas languages with ambiguous inflectional systems in expressing tense may present higher rate of RIs until the child overcomes this morphological learning problem (Kupisch and Rinke, 2007; Blom and Wijnen, 2006). Kupisch and Rinke (ibid.) determined the ambiguity of a given language by establishing those verbs in children's language, which are morphologically ambiguous between finite and non-finite, e.g. English verbs unless marked for 3rd person or past tense. Their results are calculated based on the combined total of non-finite and ambiguous verbs.

Kupisch and Rinke (2007) also analysed adult data to establish whether the presence of RIs in children's speech is the result of the presence of non-finite verbs and RIs in adults' speech. The adult data, however, exhibited very infrequent use of non-target-like infinitive forms, i.e. RIs, and therefore Kupisch and Rinke (ibid.) suggested that the verb forms in the input alone cannot account for the rate of RIs produced by the children, and therefore, it has no overall effect in the acquisition process.

5.5 Non-finite forms in *Modern Greek*

In Modern Greek all verbal forms are marked for Aspect and Agreement (Tsopanakis, 1994; Holton et al., 1997/2004). Modern Greek has no infinitives; the distinction between *finite* and *non-finite* of Romance languages operates in terms of *dependent* and *non-dependent* for tense specification (Holton et al., 1997). There is however a *non-finite* verbal form used for the formation of periphrastic (perfect) tenses, such as *perfect*, *pluperfect* and *future anterior*. For the formation of these tenses, an auxiliary verb needs

to be employed i.e. *echo* (= to have). The auxiliary verb is inflected while the *non-finite* form remains unaltered:

Perfect Tense of the verb *grapo* (= I write)

echo *grapsi* (= I have written)

echis *grapsi* (= you have written)

echi *grapsi* (= he/she/it has written)

echume *grapsi* (= we have written)

echete *grapsi* (= you have written)

echoun *grapsi* (= they have written)

The other two periphrastic tenses are formed in a similar way. The form defined *non-finite* is the past participle *grapsi* (Tsopanakis, 1994; Holton et al., 1997).

5.5.1 Verbal mood

There are three verbal Moods in MG; indicative, subjunctive and imperative. There is a fourth type referred to as *potential* or *facultative* which corresponds roughly to the conditional. Facultative mood is formed from the past indicative preceded by the particle *tha* (future), *na* (facultative), *as* (advisory, exhortative) in addition to the auxiliary for the formation of pluperfect.

5.5.2 Subjunctive and future

Subjunctive is relevant in this study. According to Holton et al. (1997), subjunctive expresses wish, desire, request, order, permission etc. from the speaker's point of view.

The subjunctive in MG is formed with the addition of the modal particles *na/as* preceding the verb. The verb with imperfective stem corresponds to the present tense subjunctive while the aorist subjunctive is formed with the perfective stem:

verb *grafa* (= write)

Present subjunctive

Aorist subjunctive

na *grafa* - 1st singular

na *grapso* - 1st singular

na *grafis* – 2nd singular

na *grapsis* – 2nd singular

na *grafi* – 3rd singular

na *grapsi* – 3rd singular

Future tense is formed in a similar way using the particle *tha*.

The use of subjunctive in main clauses expresses wish or request when used with *na* and exhortation, admonition, or consent when used with *as* (Holton et al., 1997; Triantafillidis, 2002; Klairis and Babiniotis, 2004).

Na/tha have been analysed as functional elements occupying a position within the Inflectional Phrase (Veloudis and Philippaki-Warburton, 1983; Campos, 1989; Tsimpli, 1990; Rivero, 1994 among others) or within the Complementiser Phrase domain (Agouraki, 1991; Tsoulas, 1994; Roussou; 2000).

Roussou assumes an extended CP domain along the lines of Rizzi (1997) and her C_{Mood} is roughly equivalent to Rizzi's Fin (the lowest of his CP projections). Roussou further assumes that *na* (but not *tha*) raises to a higher $C_{\text{Op(erator)}}$ position (roughly Rizzi's Force), thus accounting for differences in the relative ordering of *na* and *tha* with respect to negation, and reflecting the fact that *na* is both a modal element and a clausal typing element that can occur in a root clause. In Roussou's terms, *na* spells out both Mood and Force features.

... [C_{Op} [(Neg) [C_M *na/ta* [...I...]]]]

(Rousou, 2000)

5.5.3 *The Dependent*

The dependent is a verbal form with two properties: the perfective and non-past (Holton et al., 1997). Those properties are manifested by the use of the perfective stem and its inflection:

grapso (= write/NonPastPerfective)

diavaso (= read/NonPastPerfective)

agapiso (= love/NonPastPerfective)

Holton et al. (ibid.) suggest that the dependent does not constitute an independent tense or an independent mood but appears normally in constructions where it is preceded by one of the particles *tha*, *na* or *as*. With the future particle *tha* it forms the perfective future; with the particles *na* and *as* it forms the perfective subjunctive, referring to a complete action. According to Holton et al. (ibid.) the basic difference between the dependent and the other verb forms is that the dependent requires always the presence of one of the particles *tha*, *na* or *as*, while the other verb forms may occur either after the same three particles or independently of them.

5.5.4 Imperatives

The main function of the imperative is to express desire in the form of *order*. This is not the only one however. Imperative expresses also *advice*, *exhortation*, *pray*, etc. and *license* for the speaker to do something (Klairs and Babiniotis, 2004). Imperative in Greek forms only the 2nd singular and 2nd plural person. Imperative is tenseless, i.e. there are forms without time distinction. Klairs and Babiniotis (ibid.) suggest that imperative is by definition positioned in the future, since something is *ordered* to happen in the future (or after the end of the utterance). There are two forms of the imperative, i.e. the imperfective-present and the perfective-aorist. These two imperative forms are not related to tense:

| <i>Verb grafo</i> (= write) | 2nd Singular | 2nd Plural |
|-----------------------------|--------------------------------|------------------------------|
| Imperfective stem | <i>graf-e</i> | <i>graf-ete</i> |
| Perfective stem | <i>graps-e</i> | <i>graps-ete</i> |

The difference between the two forms is only an aspectual difference, i.e. the perfective stem expresses telicity and the imperfective stem is atelic (the action is in progress) (Klairs and Babiniotis, 2004). Imperative often does not distinguish between the two forms (used in the same context):

DjavazeIMPERF/djavasePERF perisotero kathe mera kai tha to perasis to mathima

Read/IMP more every day and you will succeed the exam

Ante, djavazeIMPERF/djavasePERF to sta grigora na figoume

Come on, read/IMP it quickly so we can go

(Adapted from Klairis and Babinotis, 2004)

Third person imperative can be expressed using the subjunctive mood and the particles *na/as*. Therefore, third person imperative is identical with the third person subjunctive and the difference can only be understood from the context or intonation (Tsopanakis, 1994).

Negative imperatives are formed in a similar way with the use of the subjunctive and the negation *min*:

2sg. imperative: *grafe* (= write/2sg/Imperative)

2sg. negative imperative: *min grafis* (= don't write/2sg/Subjunctive)

(Holton et al., 1997)

5.5.5 Person marking

Greek has a particularly rich inflectionally morphology (Klairis and Babinotis 2004; Holton et al., 1997). Below are the endings of the agreement paradigm for the two conjugations in the Active Voice:

| Person | 1 st conjugation | | 2 nd conjugation | |
|-----------------|-----------------------------|----------|-----------------------------|---------------------|
| | Singular | Plural | Singular | Plural |
| 1 st | -o | -ome/ume | -(a)o, -o | -ame/ume, - ume |
| 2 nd | -is | -ete | -as, -is | -ate, -ite |
| 3 rd | -i | -une | -a(i), -i | -an(e)/un(e), un(e) |

5.6 Linguistic accounts of RIs in Greek

In the following section a range of accounts of RI status in children's early language focusing on Modern Greek will be presented.

5.6.1 *Modality in Modern Greek (Stephany 1981, 1986, 1995)*

Stephany's (1981) study of the acquisition of Modern Greek was the first one to identify modal and non-modal verbs in early Greek.

Stephany's data, partly³⁵ available in CHILDES, comprises of four children learning Greek. The data collection consists of recordings during a 2-week period. Each recording is of 60-90 minutes duration and the age range is from 1;9 to 2;9 collectively (Stephany, 1981, 1986).

In her study, Stephany (1981) divided the analysed utterances into modal and non-modal. Stephany found that children use these categories in combination with the imperfective and perfective aspect. Children's verbal forms with perfective stem were almost always used in a modal manner (Stephany, 1981). Therefore, most modal utterances were used with the perfective stem whereas non-modal expressions were predominantly used with the imperfective stem. Stephany (ibid.) also proposed a classification into stative and dynamic verbs. Stephany (ibid.) found that children used stative verbs exclusively in non-modal utterances and mostly with the imperfective stem whereas dynamic verbs are used almost always in modal expressions and mostly with the perfective stem. Stephany (ibid.) reported 93% of stative verbs are used in a non-

³⁵ There are 23 recordings available in CHILDES.

modal expression with imperfective stem whereas 73% of resultative dynamic and 64% of non-resultative dynamic verbs are used with a modal meaning with perfective aspect.

Stephany (1986) reported that the use of modal particles *na/tha* for the formation of subjunctive and future respectively, was not reliable during the early stages of children's development. Those particles were often reduced to their vowel *-a*, or omitted:

22. (Eve 1,9) a pame sti thalasa a banjo
 PRT:a go/1pl to the sea for a swim
We will/let's go to the seaside to swim

23. (Spiros) pio vavasi
 *Spiros 0tha/na read/3sgPERF
 (= o Spiros tha/na djavasi)
Spiros is going to/wants to read

24. (Mairi) pari gugunaki May
 *0tha/na take/3sgPERF piggy Mairi
 (= i Mairi tha/na pari to gurunaki)
Mairi is going to/will take the piggy

(from Stephany, 1986)

Therefore, the distinction between subjunctive and future is not always possible. Stephany (ibid.) proposed that at this stage children's subjunctive forms are

‘plurifunctional’, and thus used to express *predictions* (future interpretation) as well as *wish* or *intention* (subjunctive interpretation).

Stephany (1981) also studied the child directed speech produced by the adults in the data collection. She found that adult’s distribution of perfective and imperfective stems across the used verbs is very similar to the one observed in children’s speech. Stephany also investigated adult directed speech. Stephany found that modals do not occur as often as in child directed speech.

Stephany in her studies did not specifically designate the term RIs in her findings of children’s speech production. She was, however, the first to identify those modal expressions in children’s early speech. Such expressions correspond to RIs. Stephany also classified those modal expressions in their semantic and inflectional context. Modal expressions were used very early in the acquisition of Greek as semantic anchors expressing concepts such as *possibility*, *necessity*, *intention* and *wish*. Stephany suggested that her findings are in disparity with Brown’s (1973) claim about the acquisition of English. According to Brown (ibid.) children acquiring English at a very early stage, i.e. below MLU 1.75 do not have tense, aspect, and mood categories in their grammar.

5.6.2 First approaches in the analysis of RIs – the -i form

Varlokosta, Vainikka and Rohrbacher (1996, 1998) (henceforth VVR) were the first to discuss and account for the use of non-finite forms in early Greek as corresponding to RIs.

VVR (ibid.) aimed to determine whether there is a corresponding RIs stage for children acquiring Modern Greek.

VVR analysed a spontaneous speech collection from the Stephany Corpus available in CHILDES database (MacWhinney, 2000; Stephany 1995). The data collection consists of recordings of three children ages between 1;9-2;5, namely, Janna, Mairi and Spiros.

According to VVR there are some verbal forms which could represent the Greek analogue of infinitive of the other child languages. These are:

- a. A verbal form combined with the “*na*” particle, which corresponds to the Greek subjunctive.
- b. A verbal form with *-i* suffix, used with both aspects, namely, either with the perfective or the imperfective stem. This form is referred to as the past participle³⁶ of periphrastic tense without the auxiliary (VVR, 1998).

For the first option VVR found in their analysed data that such constructions are often presented in main clauses, against the adult model. This fact suggested, in the first instance, the similarity of these forms with the RIs of other languages. Although the subjunctive appears to be a good candidate for RIs, VVR (ibid.) after detailed examination, present a series of arguments against identifying the *na*-construction as the equivalent of RIs.

VVR proposed that children’s use of the subjunctive particle develops in two stages. VVR claimed that the particle *na* at an early stage is not used productively and is often omitted. The proportion of *na*-clauses used during the early stages is lower than RIs (VVR, 1998). Finally, *na* emerges at the same time as *tha* (future particle), agreement, and modals. VVR suggested that this is evidence of a more advanced stage

³⁶ Although VVR proposed the participial analysis, *i-forms* are often referred to as the 3rd singular, being homophonous to the participle. In this thesis, the terms *participle-participial* will be only used when referring to VVR’s participial analysis in relation to RIs.

than the RI stage. VVR (ibid.) concluded that early *na*-clauses are the same as in the adult language. They specifically argued against identifying the RI equivalent as the child's version of the subjunctive/future structure and therefore they immediately abandon this first alternative.

The second option (used with the perfective stem) is claimed to hold the comparison with the infinitives of other languages, being a participle form, and hence less marked. VVR proposed two stages of acquisition about *-i* forms:

- **Stage 1** where the *-i* form is used in more than half of the cases. A high rate is used in a non-3sg context. Verbal forms different from 3sg are rarely used and without overgeneralisation; this is evidence for AGR. There is no evidence for Tense or Modals.
- **Stage 2** *-i* forms are used less than at the first stage and mostly in the appropriate context of 3sg. The AGR paradigm is used productively. Modals and Future Tense are acquired.

(Varlokosta, Vainikka and Rohrbacher, 1998)

From the data analysis VVR proposed that the RI candidate in child Greek is the verb form with suffix *-i*. In traditional grammars this form is ambiguous between the 3rd singular person of present tense and the active participle that with the auxiliary forms periphrastic tense. VVR claimed that children go through two stages in what regards the use of such form.

During Stage 1 (Spiros 1;9, Janna 1;11) the *-i* form is overused and mostly in a non-3sg context. VVR claim that during Stage 1 Greek children go through a

prefunctional stage, namely a stage during which their grammar lacks INFL related elements (Radford 1990, Tsimpli 1992). During this stage children only project the lower “lexical” part of the phrase (VP, NP) and the result is the production of RIs which they suggest occupy the VP (Hyams, 2002). VVR thus concluded that during Stage 1 the early grammar does not project Inflectional categories, such as AGRP and TP, in other words, that children’s grammar is prefunctional.

During Stage 2 (Janna 2;5, Mairi 1;9) *-i* forms are used less than the first stage and predominantly in 3sg context. Stage 1 is referred as the Early Non-finite Stage i.e. the RIs Stage equivalent.

VVR found that during Stage 1 Spiros’ (1;9) and Janna’s (1;11) use of *-i* forms was as high as 76% and 51% respectively. During Stage 2, however, the rate of *-i* decreased, i.e. 35% for Janna (2;5) and 38% for Mairi (1;9). Moreover, most of the *-i* forms used during Stage 1, are used incorrectly, i.e. in a non-3sg context. This is, 40% of Spiros’ and 38% of Janna’s *-i* forms are used in non-3sg context.

Child1: tuto seli

this want/3sgIMPERF

Child2: fai

Eat/3sgPERF

Child3: nitsi tola!

Open/3sgPERF now

(VVR, 1998)

Evidence for the above proposal also derived from the distribution of null and realised subjects across the corpora. VVR found that during Stage 1 the overgeneralisation of the *-i* forms, occurred mostly with null subjects (66% of correct *-i* occurred with null subjects vs. 86% of incorrect *-i* occurred with null subjects). The claim is that realised subjects occurred predominantly with correct agreement, and thus, a realised subject implied correct agreement, while incorrect agreement or non-agreement implied null subject use. This is considered to provide a further argument in support of the claim that the *-i* form is an early non-finite form equivalent to RIs in other languages.

In their paper VVR also claimed that the overuse of the *-i* form occurred mostly with the Perfective stem than with the Imperfective stem. Their findings suggested that more than half of the perfective *-i* forms involved incorrect uses of the *-i* suffix. The majority of these forms (14 out of 16) involved verbs with no morphological stem distinction (single-stem). In Spiros' data 53% of *-i* forms with the perfective stem were used incorrectly and in Janna's speech production 65% of *-i* forms with perfective stem were used in a non-3sg context.

VVR concluded that children learning Greek at their earliest stages of language acquisition used a participial form which had the same distributional characteristics of RIs in other languages. VVR proposed the use of a broader term, i.e. Early Non-finite Form. In terms of phrase structure VVR proposed that children used these forms that project as little as possible of the functional categories available in the adult language. Children at an early stage of language acquisition used the smallest tree and in this sense their functional categories system is subject to maturation (VVR, 1998).

The VVR studies provided very informative findings in support of a RI analogue stage in child Greek. Their study made a new contribution in the analysis of non-finite

forms, providing evidence for a developmental stage during which children express ‘whatever unmarked non-finite form the adult language possesses’ in a non adult-like manner (Rizzi, 1994). The analysed data, however, is the same set of data is used throughout these studies, leaving limited space for testing the proposed analyses with data that have not been analysed before.

5.6.3 *Bare Perfectives (Hyams 2002, 2005, 2007)*

Hyams’ (2002) study was a reply to VVR’s analyses on the Greek equivalent of RIs. Hyams re-analysed the data from a continuity point of view (as opposed to the maturation account proposed in VVR). She argued that the Greek data analysed in VVR did not show evidence for a pre-functional stage and that children’s speech converges early to the adult language. Hyams also proposed an analysis under the term *bare perfectives* against the participial analysis discussed in VVR. Such a ‘participial’ analysis has been the theme of a debate between VVR and Hyams in what concerns the analysis of RIs. The bare perfective is claimed (Hyams *ibid.*) to have a modal/future interpretation.

Hyams (2002, 2005) investigating the status of RIs in early Greek based on the data presented in VVR (1998), proposed that children use a form, called *bare perfective*, which in adult language, is expressed by the *na/tha* clause, i.e. the subjunctive mood, except that it lacks the modal particle.

Hyams claimed that children have essentially correct morpho-semantic mappings. The bare perfective like the RIs, typically expresses the child’s wish, need, and intention with respect to some eventuality. In other words, it has a modal or irrealis interpretation.

According to Hyams, the bare perfective shares with the RIs the following properties:

- It is non-finite, as evidenced by the lack of productive agreement.
- It has a modal or irrealis meaning, that is, it is volitional, directive, or intentional.
- It is restricted to eventive predicates.
- It co-occurs with finite clauses.

(Hyams 2005)

According to Hyams, the irrealis interpretation of the bare perfective is provided by an active MoodP. Hyams adopted Roussou's proposal (2000), who argued that *na* and *tha* are both generated in a lower C head that is specified for mood $-C_{\text{Mood}}$. According to Roussou (2000), MoodP may be licensed by a modal element merged in its head or through checking by the appropriate features in the verb. In the adult phrase structure, MoodP is licensed through merge with the presence of the modal particle. The child's structure though, is different in that *na/tha* particles are not contained in the clause.

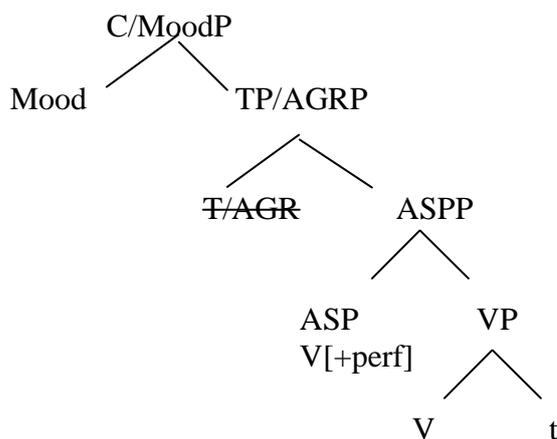
Hyams suggested that in children's grammar the phrase structure is based on an aspect-mood feature connection, combined with the unspecification of T/Agr. Hyams proposed that the aspect feature like the perfective one (FPerf) matches both, and hence, may license either Asp or Mood heads. In other words, the child grammar of Greek has the same adult clause structure, as above, and it is the aspectual feature in the bare perfective (FPerf), that is responsible for the licensing of Mood in the child's grammar.

The difference between adults and children can be summarised as to whether Mood is licensed by the merging of a modal particle *na/tha* or through checking by the

aspect feature in the verb. This explains the trade-off between bare perfective and the use of the *na/tha* particles. The productive use of modals correlates with the use of both perfective and imperfective verbs in modals contexts. Finally the lack of agreement on the bare perfective follows as a direct structural consequence of the aspect-mood relation discussed above, i.e. that T/AGR layers need to be unspecified in order for the Perfective feature to licence the MoodP, according to the locality condition (Chomsky 1995).

The mechanism adopted in the following structural representation proposed in Hyams, is a checking mechanism, namely Attract F that incorporates this locality condition (the Minimal Link Condition). In the early grammar of Greek, Mood attracts the [+perfective] feature in the verb, and is in that way licensed.

In the phrase structure, it follows from the locality condition that there can be no features intervening between Mood and the perfective verb. In other words, T/AGR must be unspecified:



(Hyams, 2002)

Given the structural constraints on Attract/feature checking, the bare perfective is necessarily a non-agreeing form, with the unmarked *-i* affix emerging as the default

where AGR is unspecified. According to this analysis, the availability of bare perfectives and RIs in early grammar is thus plausibly related to the lack of modals and not to any specific differences between adult and child grammars.

Hyams discussed a variety of languages both with RI effects and without. In all cases the data used in this study is data analysed in other studies. So, for child Greek, Hyams' data is sourced in CHILDES database (MacWhinney, 2000; Stephany 1995), i.e. same data analysed by VVR.

Hyams' analysis showed that AGR is used correctly in children's early language and that AGR features are productive:

| | <i>Stage</i> | <i>% of correct AGR use (1SG and 2SG)</i> |
|--------|--------------|---|
| Janna | 1 | 91% |
| Spiros | 1 | 83% |
| Janna | 2 | 99% |
| Mairi | 2 | 100% |

(Hyams, 2002, p. 11)

Based on these results, Hyams claimed that agreement is productive and the two stages proposed in VVR are not confirmed empirically.

Hyams also proposed that perfective *-i* forms (3SG) are also used with incorrect agreement more often than the other persons:

| | <i>Stage</i> | <i>% of correct AGR use (-i forms)</i> |
|--------|--------------|--|
| Janna | 1 | 62% |
| Spiros | 1 | 60% |
| Janna | 2 | 100% |
| Mairi | 2 | 70% |

(Hyams, 2002, p. 13)

In summary, Hyams claimed that there is not enough evidence to propose two stages in the acquisition of Greek, where during Stage 1 children have not developed yet inflection related categories. Instead, she claimed that children learning Greek have the same clause structure as adults. Early Greek bare perfective is used to express the irrealis interpretation replacing adult modals, i.e. satisfying the grammatical function of adult *na/tha* clauses. According to Hyams, the bare perfective is positioned in MoodP and there is a trade-off between this early child form and *na/tha* adult modal clauses.

In Hyams (2007) the temporal and modal meanings associated with RIs are examined. Hyams (ibid.) claimed that in non-finite constructions where tense is not specified, the event organisation of the predicate determines the temporal reference of the construction. Such event structures are defined by the predicates' aspectual properties. Hyams' study analysed a variety of early languages such as Dutch, English, Russian and Greek for their temporal and modal characteristics of non-finite clauses.

Building on the Punctuality Constraint (Giorgi and Pianesi, 1997), a closed event cannot be simultaneous with the utterance time (speech event). Building on the Punctuality Constraint, Hyams (ibid.) proposed the *Closed Event Hypothesis* according

to which in children's non-finite clauses a temporal reference is allocated according to the topological property of event closure. The Closed Event Hypothesis is claimed to have two requirements (Hyams, 2007); the *Temporal Anchoring Requirement* (TAR) according to which events must have a temporal reference ordered with respect to the utterance time and the *Default Anchoring Requirement* (DAR) according to which if tense is unspecified then the event time coincides with the utterance time.

The Punctuality Constraint together with the Closed Event Hypothesis and the Temporal and Default Anchoring Requirements lead to a set of options for the expression of temporal properties in children's non-finite clauses, named by Hyams (ibid.) *Linking Algorithm*:

Linking algorithm

- (a) If open, link e (= event) to UT (= Utterance Time)
- (b) If closed/perfective, then
 - (i) if sectional restrictions permit, insert null modal, and link to UT,
 - (ii) if not, insert telos (telic predicate), and link to UT.

(Hyams, 2007: 256)

Option (a) is an open event where the event is ongoing at the utterance time. If the event is closed (b) then there are two options: (i) a null modal is inserted functioning as a temporal anchoring to the utterance time (since the closed event is not accessible to the utterance time) and (ii) a second event (telos) is inserted as a temporal anchor to the utterance time (telic anchoring).

According to the *Linking Algorithm*, Hyams (ibid.) suggested that the Greek Bare Perfective is a closed event due to the perfectivity of the verb; the insertion of a

null modal (b-i) satisfies the anchoring requirement. To explain the fact that Greek speaking children produce both telic (where the anchoring to UT can be provided by the telos event) and atelic bare perfectives, Hyams (ibid.) suggested that the two options are not equally accessible (one quarter of all BPs are atelic) and that a telic event is only inserted when the null modal insertion is blocked. In finite clauses tense is specified and this provides the anchor to UT.

5.6.4 *New evidence for the non-finite forms (Varlokosta 2002, 2005)*

In Varlokosta (2002) there is some new evidence for the correlation between *-i* form and RIs. Varlokosta claimed that such new evidence provides further support to the original proposal, postulated in 1996/1998, that *-i* forms in early Greek equals RIs in other languages, i.e. they are the Greek RI analogue.

Varlokosta proposed that there are three phenomena in children's grammar that are directly related to the RI stage in Greek and they constitute evidence for such a stage. Those are:

- Clitic omission with the use of non-finite forms with perfective aspect with the suffix *-i*
- Reduced use of past tense during the RI stage
- Determiner drop occurring with the use of non-finite forms with perfective aspect with the suffix *-i*

The first evidence is the correlation between the use of infinitival forms and the omission of clitics. Based on the coding in the Stephany corpus, Varlokosta found 31 cases of clitic object omission in the speech of Spiros (out of 33 contexts with an

obligatory object or clitic) and only 6 in the speech of Janna (out of 7 contexts with an obligatory object or clitic) both children at Stage 1. For Greek, the correlation holds between the use of verb forms with perfective aspect with the suffix *-i* and clitic object omission:

Adult: afti ine i kardula

(Spiros 1;9)

This is the heart/DIM

This is the little heart

Child: kepasi i gazula a fai o likos

Cover/3sgPERF* the*NOM heartACC/DIM SUBJ/FUTparticle eat/3sgPERF
the wolfNOM

I (will/am going to) cover the little heart for the wolf not to eat (it)

Adult: etsi?

(Janna, 1;11)

Like-this?

Like this?

Child: nitsi tola!

Open/3sgPERF* now

(You) open (it) now

Child: anitsi!

Open/3sgPERF*

(You) open (it)

(Varlokosta, 2002, p. 144)

Second evidence is the poor tense verbal paradigm in this stage. What is suggested is that children do not use [+past] verbs in their speech production during the RI stage, and there are very few occurrences of [+past] verbs observed. Varlokosta (2002) reported that Janna (1;11) produced seven verbs in the past tense but they are all instances of the verb *kano* (=make) while Spiros (1;9) produced five past tense verbs out of 127:

Ta-vale (= s/he put them)

Epese (= it fell)

Bike (= it got in)

Fonakse (= s/he screamed)

(Varlokosta, 2002, p. 142)

Finally, the third evidence suggesting the similarity between *-i* forms and RIs, is the correlation between determiner drop and use of RIs. Based on Marinis (1998), Varlokosta suggested that all instances of determiner drop occur in this context, namely *-i* forms and no-agreement.

This new evidence will be analysed³⁷ and discussed further in the second part of this chapter, in Maria's and Eve's corpora.

³⁷ An analysis of the third piece of evidence was not viable because of the lack of coding related to DPs, being beyond the scope of the current thesis.

5.6.5 Imperatives – RI analogue in Null Subject Languages

Salustri and Hyams (2003) proposed that the RI analogue in null subject languages is the imperative. According to their proposal, RIs are identified by three temporal properties:

- a. RIs are tenseless verbs in root contexts
- b. RIs typically have a modal or irrealis meaning expressing volition, intention, or direction with respect to some eventuality [MRE= Modal Reference Effect]
- c. RIs are typically eventive [EC= Eventivity Constraint]

(Salustri and Hyams, 2003)

The *bare perfective* (BP) is a non-finite, non-agreeing form. According to Salustri and Hyams (ibid.) Greek speaking 2-year olds use the BP, which is the RI equivalent. Building on this proposal, Salustri and Hyams suggested that imperatives are irrealis, i.e. they express a direction and they are restricted to eventive predicates; and they are tenseless.

Cross-linguistic variation is explained in terms of selection, i.e. RI languages choose an infinitive while null-subject languages choose the imperative. Those two options is claimed to have similar structural and licensing properties despite using different non-finite forms. Building on Han's description of imperative as a form marked with an 'irrealis' feature (Han, 2001), Salustri and Hyams claimed that children begin using imperatives before they reach two years of age. Imperatives, as opposed to RIs and BP, are fully grammatical when used in children language. Their analysis of

Imperatives is mainly proposed for early Italian. For early Greek it is the perfective feature of the verb marked with the irrealis meaning (Salustri and Hyams, 2003).

Salustri and Hyams proposed the Imperative Analogue Hypothesis (IAH) according to which:

- a. In null subject languages imperatives will occur significantly more often in child language than in adult language
- b. In child language imperatives will occur significantly more often in null subject languages than in RIs languages.

(Salustri and Hyams, 2003: 4)

The rationale behind this hypothesis is that over time some imperatives will be replaced by modals, just as RIs and BPs trade off with modals in Dutch and Greek.

In Salustri and Hyams (2006) the Imperative Analogue Hypothesis is further examined using data of children speaking other languages, i.e. Italian and German, Dutch (Blom, 2002), Icelandic (Sigurjónsdóttir, 2005), Catalan (Bel, 2001 cited in Salustri and Hyams, 2006), Spanish (Arbisi-Kelm, 2005 cited in Salustri and Hyams, 2006), Hungarian (Londe, 2004 cited in Salustri and Hyams, 2006) and Slovenian (Rus, 2004 cited in Salustri and Hyams, 2006).

Salustri and Hyams (*ibid.*) predicted that imperatives will occur more frequently in children's language than in the adults'. Their analyses suggested that this is indeed the case, only for null-subject languages. Therefore, they claimed, that in languages with a strong RIs effect (e.g. German, Dutch, Icelandic) imperatives are not frequent in children's speech, whereas in null-subject languages (non-RI languages such as Italian,

Spanish, Catalan, Hungarian and Slovenian) the use of imperatives is very high at the early stages of speech production and decreases with time.

5.6.6 Semantic interpretation and modal properties of RIs and BPs

RIs often received a modal interpretation, which refers to a likely future event. It has also been claimed that their temporal reference is essentially free (Poeppl and Wexler, 1993; Hoekstra and Hyams, 1998; Hyams, 2002/2005). The majority of RIs constructions are eventive predicates, whereas finite constructions allow both eventive and stative predicates. Studies in Dutch (Wijnen, 1997; Jordens, 1990) and in French (Ferdinand, 1995) observed that the majority of RIs examined in their data occurred with eventive predicates and had a temporal interpretation.

Hoekstra and Hyams (1998) building on these analyses of RIs future interpretation, claimed that in the examined data in various languages, RIs have a temporal reference that entails a modal irrealis interpretation. Such interpretation conveys messages of deontic and boulemaic modality, expressing necessities and desires (Hoekstra and Hyams, 1998: 91). Hoekstra and Hyams (ibid.) therefore formulated the Eventivity Constraint, according to which RIs are restricted to event-denoting predicates. Based on cross-language analyses of children's early speech presenting a RI effect (De Haan, 1986 and Wijnen, 1997 for Dutch; Ferdinand, 1995 for French; Van Gelderen and Van der Meulen, 1998 for Russian; Ingram and Thompson, 1996 for Swedish; Meisel, 1990 for German; to mention but a few) Hoekstra and Hyams (ibid.) also formulated the Modal Reference Effect (MRE), according to which RIs in children's data have predominantly modal reference.

Hyams (2002) claimed that the RI stage in children's language is not optional but is the result of children's attempts to express the irrealis/realis opposition.

According to the Semantic Opposition Hierarchy, the category of Mood and in particular the realis-irrealis opposition is the most primitive one (RIs are irrealis forms whereas finite forms are realis). Irrealis expresses desire, necessity or future. According to Hyams (ibid.) children use non-finite structures in their attempt to map specific semantic features onto specific morphosyntactic structures.

Giannakidou (2002) suggested that Greek verb forms are unambiguously eventive or stative, depending on whether they have perfective or imperfective aspect. Thus, perfective verb forms are always eventive. Building on Giannakidou, Varlokosta (2002, 2005) reported that a proportion 87% of *-i* forms with no agreement involves eventive predicates. Varlokosta proposed that there is a future or modal interpretation involved in the majority of *-i* forms, and thus, there is a restriction regarding their temporal reference, i.e. the modal interpretation. The eventivity constraint as well as their modal/irrealis interpretation showed that these forms have parallel semantic interpretation to RIs in other child languages. It is claimed that this evidence provides further support for the non-finite status of these forms.

Hyams (2005) identified the relation between perfectivity and modality with regard to the BPs. The relationship between modality, perfectivity and eventivity is as follows:

deontic modality > perfectivity > eventivity

(Hyams, 2005)

Given the association between deontic modality and perfectivity and eventivity, it follows that the bare perfective will be restricted to eventive predicates according to the eventivity constraint discussed above.

Christofidou and Stephany (2003) in their study of acquisition of verb morphology analysed spontaneous data of two children aged 1;7 to 2;1. Their results suggested that atelic predicates (durative activities and states) are mainly used with present tense whereas telic predicates are mainly used with past perfective with a resultative meaning.

Stephany and Voeikova (2003) suggested that in children's language (and also in child directed speech) there is a strong two-way dependence between perfective and imperfective aspect and state and action verbs from one standpoint, and perfective and imperfective aspect and telic and atelic verbs from another standpoint. In their analysis of Christos' data collection, Stephany and Voeikova proposed that stative verbs are exclusively used with imperfective aspect, telic verbs mostly used with perfective aspect (75%) and atelic verbs with both perfective and imperfective aspects (50%-50%). Varlokosta (2002, 2005) reported that a proportion 87% of *-i* forms with no agreement involves eventive predicates.

5.6.7 Person and Infinitive features and RI stage length continuum (Perales, Licerias and Bel)

Perales et al. (2006) and Licerias et al. (2006) proposed a RI length *continuum* of languages based on the status of Person [+/-P] and Infinitive [+/-R] features in each language.

Building on Hyams' (2001) *Semantic Opposition Hypothesis*, both studies departed from the assumption that mood is the most primitive semantic opposition (realis vs. irrealis) and that children map these primitive meanings into inflectional constituents. Realis expresses ongoing or actual events (both in the past and present) whereas irrealis expresses desire, necessity or futurity of an event.

Perales et al. (ibid.) used data from Spanish (Liceras et al., 1999), Basque (Ezeizabarrena, 2002) and Catalan (Bel, 2001). Their analyses found that in all data, children used RIs to convey both realis and irrealis values. Their data analyses concluded that a) there is a RI stage across all languages although limited in null subject languages and b) in null subject languages RIs may express both an irrealis and a realis reading (contrary to non null subject languages).

Perales et al. (ibid.) proposed that Person and Infinitive features are responsible for the realis/irrealis values manifested in children's language. [+/-P] and [+/-R] determine the realisation of the irrealis/realis semantic opposition (Perales et al., 2006). Therefore, realis is expressed with the use of inflected forms whereas irrealis is realised with the use of bare forms. According to Perales et al. (ibid.), in a given language, all inflected forms with person marking [+P] will realise the realis interpretation, whereas, bare forms [-P] express the irrealis meaning. In languages with infinitives [+R], Perales et al. (ibid.) suggested that this competes with the person marker in expressing the irrealis.

Table 5.1 below, adapted from Perales et al. (2006), illustrates how [+/-P] and [+/-R] features determine the realis/irrealis opposition in different languages:

Table 5.1: Realis and irrealis opposition expressed in [+/-P] and [+/-R] features in different languages.

| | | | Realis | Irrealis |
|------------------|---------------|---------------|----------------------------|--|
| | [+/-P] | [+/-R] | | |
| Dutch, German | - | + | Inflected | Infinitives |
| Greek | + | - | Inflected | Bare subjunctive/perfective |
| Italian, Spanish | + | + | Inflected Infinitives | Bare 3 rd person Infinitives |
| English | - | - | Inflected Bare forms | Semi-auxiliaries Bare forms |

(Perales et al., 2006, p. 95)

Liceras et al. (2006) used the same data used in Perales et al. (2006). Building on Rizzi's (1994) Truncation Hypothesis, the study assumed RIs to be VPs in a tree diagram, and according to the VP Internal Subject Hypothesis, it is assumed that DP subjects are generated in the VP. This is similar to VVR's proposal.

Liceras et al. (ibid.) also proposed that the RI stage length is a *continuum* subject to the distribution of Person [+/-P] and Infinitive [+/-R] features. Languages with [+P] and [+R] such as Catalan, Italian and Spanish are suggested to have a short and less obvious RI stage than languages with [-P] such as Dutch and German (Liceras et al., 2006). In English with [-P] and [-R] features the RI stage is proposed to be even longer (Liceras et al., ibid.). Figure 5.1, adapted from Liceras et al. (ibid.) illustrates the *continuum* of RI stage length determined by the combination of [+/-P] and [+/-R] features:

What is clear from all the above, is that children learning Greek use a non-adult verbal construction. This phenomenon is less extensive when compared to non null-subject languages, such as Dutch, German etc. What is not clear yet is how widely these constructions are used in children's speech. We also do not know if this phenomenon is due to lack of some syntactic or morphological component in children's language or simply lack of input from their environment. This study is motivated by the need of analysing more data of spontaneous speech in early Greek.

5.8 Analyses

In the following section the following analyses were performed on the data of Maria and Eve:

- An error analysis was completed to check the errors performed to Maria's and Eve's speech that might be related to non-finite forms, e.g. missing modal particles.
- A modal particles analysis was carried out in order to establish the rate of use of subjunctive and future constructions and the rate of omission of such particles in obligatory contexts.
- An analysis of overuse of the *-i* forms (*-i* form 3sg and other 3sg) was performed to test the existence of two stages where the overuse of *-i* forms decreases with time.
- An analysis of agreement of *-i* form was conducted in order to identify those *-i* forms used in non-3sg context.
- An analysis of null subjects in *-i* forms was carried out to establish whether *-i* forms occur with null or realised subject.
- A stem analysis was performed to check if *-i* forms are used more with the perfective stem rather than the imperfective one.

- Analyses of Past tense and Clitic Object Omission in relation to *-i* forms was completed to test previous empirical data.

5.8.1 Errors

The following section examines the distribution of the various errors in Maria's and Eve's corpora in comparison to previous studies.

Tables 5.2 and 5.3 show the overall error rates, i.e. correct and incorrect forms, in Maria's and Eve's data respectively. The tables contain all error categories. These error categories are discussed in detail in the following section.

Table 5.2: Maria's overall error rate (verbal domain)

| Maria/Age | MLU | Correct verbs (in utterances) | Incorrect verbs (in utterances) |
|------------------|------------|--|--|
| 2;0 | 3.8 | 81.4% (n=35) | 18.6% (n=8) |
| 2;2 | 3.4 | 93% (n=80) | 7% (n=6) |
| 2;3 | 3.8 | 94.4% (n=136) | 5.6% (n=8) |
| 2;5,4 | 4.2 | 92.4% (n=147) | 7.6% (n=12) |
| 2;5,24 | 3.6 | 96.8% (n=122) | 3.2% (n=4) |
| 2;7 | 3.7 | 96.8% (n=92) | 3.2% (n=3) |
| 2;8,3 | 3.6 | 91.4% (n=64) | 8.6% (n=6) |
| 2;8,27 | 4 | 98.3% (n=118) | 1.7% (n=2) |
| Total | | 94.2% (n=794) | 5.8% (n=49) |

In Table 5.2 the overall verbal related error rate in Maria's utterances is 5.8% of all the analysed data. Maria's error rate at the age of 2;0 is at its highest i.e. 18%. After that the rate decreases over time.

Table 5.3: Eve's overall error rate (verbal domain)

| Eve/Age | MLU | Correct verbs (in utterances) | Incorrect verbs (in utterances) |
|----------------|------------|--|--|
| 1;7 | 2 | 90% (n=72) | 10% (n=794) |
| 1;8 | 1.6 | 88.2% (n=30) | 11.8% (n=4) |
| 1;9 | 2 | 94% (n=173) | 6% (n=11) |
| 1;10 | 1.9 | 94.1% (n=272) | 5.9% (n=17) |
| 1;11 | 2.3 | 98.9% (n=196) | 1.1% (n=2) |
| 2;0 | 2.4 | 96.5% (n=394) | 3.5% (n=14) |
| 2;2 | 2.5 | 97.8% (n=267) | 2.2 (n=6) |
| 2;3 | 2.3 | 99.6% (n=281) | .4 (n=1) |
| 2;5 | 2.8 | 98.4% (n=327) | 1.6% (n=5) |
| 2;6 | 2.5 | 99% (n=209) | 1% (n=2) |
| 2;9 | 2.8 | 100% (n=71) | 0% (n=0) |
| 2;10 | 2.7 | 98.3% (n=119) | .7% (n=2) |
| 2;11 | 4.4 | 100% (n=27) | 0% (n=0) |
| Total | | 97.1% (n=2438) | 2.9% (n=72) |

In Table 5.3 Eve's error rate is 2.9%. At the age of 1;7 and 1;8 Eve's error rate is as high as 10% and 11% respectively. Thereafter, the rates gradually decrease over time.

Tables 5.2 and 5.3 contain all categories of errors related to the verb in Maria's and Eve's data. Previous studies don't include all error types in their reports for comparison with the current findings. Some studies, however, provide reports on specific error type, e.g. agreement, which can be used for comparison.

VVR reported agreement errors for Janna in Stage I (age 1;11) and Stage II (age 2;5). Table 5.4, adapted from Tables 6 and 7 from VVR (1998) present Janna's overall agreement errors. Janna was selected on the basis that her speech data covers both stages as defined in VVR (*ibid.*).

Table 5.4: Janna's correct and incorrect use of agreement

| | Correct | Incorrect |
|------------|----------------------|--------------------|
| Janna 1;11 | 72% (n=49) | 28% (n=19) |
| Janna 2;5 | 99.5% (n=167) | 0.5% (n=1) |
| Total | 91.5% (n=216) | 8.5% (n=20) |

(Adapted from VVR, 1998, p. 10)

Table 5.4 shows that Janna's overall agreement error rate is 8.5%. This is much higher during the early stages of her speech production i.e. 28%³⁸ (Stage I) and decreases to 0.5% thereafter during a more advanced stage of speech production. Maria's and Eve's overall error rate is relatively low in comparison.

³⁸ The majority of agreement errors involve 3rd singular in non 3rd singular contexts. This is discussed in details in a following section.

For further comparison, Table 5.5 presents the agreement error rates in other languages (Hoekstra and Hyams, 1998).

Table 5.5: Rate of agreement error in children's languages

| Child | Language | Age | n | %error | Source |
|-----------|----------|----------|------|--------|--------------------------|
| Simone | German | 1;7-2;8 | 1732 | 1% | Clahsen and Penke 1992 |
| Martina | Italian | 1;8-2;7 | 478 | 1.6% | Guasti 1994 |
| Diana | Italian | 1;10-2;6 | 610 | 1.5% | Guasti 1994 |
| Guglielmo | Italian | 2;2-2;7 | 201 | 3.3% | Guasti 1994 |
| Claudia | Italian | 1;4-2;4 | 1410 | 3% | Pizzuto and Caselli 1992 |
| Francesco | Italian | 1;5-2;10 | 1264 | 2% | Pizzuto and Caselli 1992 |
| Marco | Italian | 1;5-3;0 | 415 | 4% | Pizzuto and Caselli 1992 |
| Marti | Cat/Span | 1;9-2;5 | 178 | 0.56% | Torrens 1995 |
| Josep | Cat/Span | 1;9-2;6 | 136 | 3% | Torrens 1995 |
| Gisela | Catalan | 1;10-2;6 | 81 | 1.2% | Torrens 1995 |
| Guillem | Catalan | 1;9-2;6 | 129 | 2.3% | Torrens 1995 |

(Adapted from Hoekstra and Hyams, 1998, p. 84)

Maria's and Eve's overall error rate (containing more error types) appears to be at a similar level compared to children's agreement error in other languages. However, there seems to be also a higher error rate also in Eve 1;7 and 1;8 vs. all other recordings and in Maria 2;0 vs. all other recordings. These high error rates drop in the following recordings. This can be interpreted as a developmental early stage in regards to the rate of error in Maria's and Eve's data compared to children's error rates from other

languages. In fact, in those early recordings the error rate is higher than the rates of other languages presented in Table 5.5, however, somewhere between the rates of Janna in 1;11 and 2;5.

5.8.1.1 *Error classification*

Errors were divided in five main categories:

A. *Missing particle*: modal particles *na/tha* used for the formation of the subjunctive and future respectively are omitted before the verb when the particle is required in the adult language. This category contains missing particles from 1st, 2nd, and 3rd singular and plural verbs.

25. *0prt to valo mesa sto cheraki
_ it put/1sgDEP into the handDIM
I will/am going to put it into the little hand

(Maria, 2;7)

26. *tola 0prt to malosso
now _ it scold/1sg
I will/am going to scold it

(Eve, 1;4)

B. *Root Infinitives*: verbal forms of 3rd singular *-i* suffix verbs with perfective form which usually involve null subject, non-agreement (checked from the context) and missing particle.

27. *0prt vali tiiorasi

_ turn-on/3sgDEP telly

he/she (wants/is going) to turn the telly on

(Maria, 2;2)

28. *0prt feri (so)kolata ti Evis

_ bring/3sgDEP chocolate to Eve

He/she (will/is going) to bring Eve a chocolate bar

(Eve, 1;10)

C. *Agreement errors*: the subject is not in agreement with the verb. In cases of null subject the agreement was checked through the context e.g. the child is referring to herself in 3rd singular person.

29. *a(f)tos echo

heNOM have/1sg

He I-have

(Maria, 2;3)

30. *i Eva patao

the Eve press/1sg

Eve I-press

(Eve, 1;9)

D. *Missing clitic*: an obligatory clitic is omitted before or after the verb.

31. *0prt 0cl kuboso ego

_ _ button/1sgDEP I

I will/am going to button it

(Maria, 2;2)

32. *0cl vazis eki?

_ put/2sgDEP there?

(could) you put it there?

(Eve, 2;0)

E. *Other*³⁹: this category contains a variety of unclassified errors, such as phonological errors, omitted verbs from an utterance, or errors that affect the grammaticality of the sentence due to wrong word order.

33. *kselis 0wh ekane mia fora?

know/2sg _(ti) did/3sg one time?

Do you know what he did once?

(Maria, 2;5,4)

34. *thelo na 0v me

want/1sg toSUBJ _ me

I want you to _ me

(Eve, 2;6)

³⁹ In Eve's data the computerised coding was designed to contain two additional error categories, i.e. phonological errors in the verb and missing verb. These two categories were merged into the category *other* for convenience of comparison with Maria's error coding.

Tables 5.6 and 5.7 present the distribution of the various errors divided by category.

Table 5.6: Maria's errors (by category)

| Maria/Age | MLU | Missing particle | RI | Agreement | Missing clitic | Other | Total |
|------------------|------------|-------------------------|-----------|------------------|-----------------------|--------------|--------------|
| 2;0 | 3.8 | 6 | 0 | 3 | 1 | 1 | 11 |
| 2;2 | 3.4 | 4 | 1 | 1 | 2 | 1 | 9 |
| 2;3 | 3.8 | 4 | 1 | 5 | 2 | 0 | 12 |
| 2;5,4 | 4.2 | 5 | 1 | 1 | 2 | 7 | 16 |
| 2;5,24 | 3.6 | 4 | 0 | 0 | 0 | 0 | 4 |
| 2;7 | 3.7 | 3 | 0 | 1 | 0 | 0 | 4 |
| 2;8,3 | 3.6 | 0 | 0 | 3 | 0 | 1 | 4 |
| 2;8,27 | 4 | 0 | 0 | 1 | 0 | 1 | 2 |
| Total | | 26 | 3 | 15 | 7 | 11 | 62 |

Table 5.7: Eve's errors (by category)

| Eve/Age | MLU | Missing particle | RI | Agreement | Missing clitic | Other | Total |
|----------------|------------|-------------------------|-----------|------------------|-----------------------|--------------|--------------|
| 1;7 | 2 | 1 | | | 1 | 7 | 9 |
| 1;8 | 1.6 | 1 | 1 | 2 | | 2 | 6 |
| 1;9 | 2 | 3 | 3 | 2 | 2 | 2 | 12 |
| 1;10 | 1.9 | 8 | 4 | 3 | 4 | 2 | 21 |
| 1;11 | 2.3 | 2 | | | 1 | | 3 |
| 2;0 | 2.4 | 14 | 6 | 6 | 1 | 4 | 31 |
| 2;2 | 2.5 | 4 | | | 2 | 3 | 9 |
| 2;3 | 2.3 | 1 | | | | | 1 |
| 2;5 | 2.8 | 3 | 3 | | 1 | | 7 |
| 2;6 | 2.5 | 2 | | | | 2 | 4 |
| 2;9 | 2.8 | | | | | 1 | 1 |
| 2;10 | 2.7 | 1 | | 1 | | | 2 |
| 2;11 | 4.4 | | | | | | |
| Total | | 40 | 17 | 14 | 12 | 23 | 106 |

It is important to point out that the error distribution presented in Table 5.6 and 5.7 does not reflect the numbers in the overall error rate in Tables 5.2 and 5.3. This is because there are often overlaps, i.e. more errors correspond to the same verbal form. Tables 5.2 and 5.3 present the verbs involved in errors whereas Tables 5.6 and 5.7 present the various types of errors (often occurred with the same verb), for example, in both Maria's and Eve's data the presence of Root Infinitives errors is linked to the omission of particles, namely root infinitives always involve a missing particle (*na*). There is also an obvious relation between agreement errors and the presence of root infinitives. These patterns are examined in detail in the following section. Clitic

omission is present in Maria's and Eve's speech production. This will also be examined in detail in the following analyses.

5.8.2 *Modal particles (na/ta) analysis*

Hyams (2002) proposed that children use a form, called *bare perfective*, which in adult language, is expressed by the *na/ta* clause, i.e. the subjunctive and future mood respectively, except that it lacks the modal particle. Hyams (ibid.) also claimed that the bare perfective like the RIs, typically expresses the child's wish, need, and intention with respect to some eventuality and therefore its meaning is very close to what in adult language is expressed with *na/ta* constructions. In other words, it has a modal or irrealis interpretation. According to VVR (1998) *na*-constructions are used less frequently during stage I, however, the subjunctive particle *na* is used productively at stage II. The rate of the use of modal particles increases as the rate of bare perfective decreases (VVR, ibid.).

Building on these claims, the following analysis will investigate the status of modal particles in Maria's and Eve's early speech.

Table 5.8 presents *na*-clauses with modal particle used and omitted in Maria's data.

Table 5.8: Maria's na-constructions with modal particle used and omitted

| Maria Age | MLU | na + V | 0na + V |
|----------------------|------------|--------------------|------------------|
| 2;0 | 3.8 | 76.9% (n=10) | 23.1% (n=3) |
| 2;2 | 3.4 | 90% (n=29) | 10% (n=3) |
| 2;3 | 3.8 | 96% (n=26) | 4% (n=1) |
| 2;5,4 | 4.2 | 97.5% (n=39) | 2.5% (n=1) |
| 2;5,24 | 3.6 | 89.4% (n=17) | 10.6% (n=2) |
| 2;7 | 3.7 | 89.4% (n=17) | 10.6% (n=2) |
| 2;8,3 | 3.6 | 100% (n=14) | 0% (n=0) |
| 2;8,27 | 4 | 100% (n=22) | 0% (n=0) |
| Total | | 93% (n=174) | 7% (n=12) |

At the age of 2;0 Maria's *na* omission is at its highest, i.e. 23% of all the na-constructions used during this recording. At 2;2 the omission of particle *na* falls to 10% and therefore at the age of 2;3 and 2;5,4 the percentage of *na* omission falls even lower, namely 4% and 2.5% respectively. At the ages of 2;5,24 and 2;7 the percentage of omitted *na* rises to more than 10% and then totally disappears after that. Maria's percentage of omission of modal particle *na* is 7% of the total of na-constructions. Maria uses na-constructions amply in her speech, i.e. 21% out of all verbal constructions in her data and mostly correctly.

Table 5.9 presents na-clauses with modal particle used and omitted in Eve's data.

Table 5.9: Eve's *na*-constructions with modal particle used and omitted

| Eve Age | MLU | na + V | 0na + V |
|--------------------|------------|----------------------|-------------------|
| 1;7 | 2 | 100% (n=1) | 0% (n=0) |
| 1;8 | 1.6 | 100% (n=0) | 0% (n=0) |
| 1;9 | 2 | 50% (n=2) | 50% (n=2) |
| 1;10 | 1.9 | 85.7% (n=12) | 14.3% (n=2) |
| 1;11 | 2.3 | 100% (n=12) | 0% (n=0) |
| 2;0 | 2.4 | 100% (n=82) | 0% (n=0) |
| 2;2 | 2.5 | 98% (n=53) | 2% (n=1) |
| 2;3 | 2.3 | 100% (n=26) | 0% (n=0) |
| 2;5 | 2.8 | 100% (n=49) | 0% (n=0) |
| 2;6 | 2.5 | 100% (n=44) | 0% (n=0) |
| 2;9 | 2.8 | 100% (n=13) | 0% (n=0) |
| 2;10 | 2.7 | 100% (n=17) | 0% (n=0) |
| 2;11 | 4.4 | 100% (n=5) | 0% (n=0) |
| Total | | 98.4% (n=316) | 1.6% (n=5) |

Eve's overall rate of *na* omission is much lower than Maria's at 1.6%. At the ages of 1;7 and 1;8 there's no omission of the modal particle *na*, however, the use of *na*-constructions is very limited too. At the age of 1;9 there's a 50% omission of the particle *na*. There are, however, very few instances and thus this is not informative. Subsequently, the use of *na*-constructions increases with age and omissions of *na* particle are very few i.e. just 1/54 at the age of 2;2 and none thereafter. Eve's use of *na*-construction out of all verbal clauses is 12.7%.

Examples of *na* used and omitted in a clause are given below:

35. selo na chorepso ego apo ki
want/1sg naSUBJ dance/1sgDEP I(emphatic) from there
I want to dance on that side

(Maria, 2;2)

36. 0na to k(l)isume
0naSUBJ it turn-off/1plDEP
We (want/are going) to turn it off

(Maria, 2;0)

37. (th)elo na to kuposo
Want/1sg naSUBJ it fasten/1sgDEP
I want to button it

(Eve, 1;10)

38. thelo to xxx xxx 0na taiso tin Agela
want/1sg the xxx xxx 0naSUBJ feed/1sgDEP the Angela
I want xxx xxx to feed Angela

(Eve, 2;2)

VVR (1998) reported that Spiros and Janna at stage 1 use *na*-clauses 0% and 8% respectively, whereas Janna and Mairi at stage 2 use *na*-clauses productively, i.e. 23%

and 17% respectively. Tables 5.10 and 5.11 present Maria's and Eve's use on *na*-clauses in their entire speech production:

Table 5.10: Maria's use of *na*-clauses

| MLU stages | Use of <i>na</i>-clauses (entire dataset) |
|----------------------------|--|
| IV (2;0 – 2;2) | 35.1% (n= 45/128) |
| V (2;3 – 2;5,4) | 22.1% (n= 67/302) |
| IV (2;5,24 – 2;8,3) | 17.6% (n= 52/294) |
| V (2;8,27) | 18.4% (n= 22/119) |

Table 5.11: Eve's use of *na*-clauses

| MLU stages | Use of <i>na</i>-clauses (entire dataset) |
|-------------------------|--|
| II (1;7 – 2;0) | 9.4% (n= 111/1193) |
| III (2;2 – 2;10) | 15.7% (n= 203/1290) |
| V (2;11) | 18.5% (n= 5/27) |

In Table 5.10, Maria's use of *na*-clauses is 35.1% during stage IV between 2;0 and 2;2, then 22.1% at stage V between 2;3 and 2;5,4. Consequently, between 2;5,24 and 2;8,3 the use of is 17.6%, and finally, at 2;8,27 is 18.4%. These rates are much higher compared to stage I reported in VVR, very similar, however, to the rates reported for stage II.

In Table 5.11, Eve's at stage II, between 1;7 and 2;0 use of *na*-clauses is 9.4%, and 15.7% during stage III, between 2;2 and 2;10. Finally, during stage V, at 2;11 the

rate of use of *na*-clauses is 18.5%. These rates correspond to the rates reported in VVR for stage II. Maria's overall rate of *na*-clauses at 21% is comparable to Janna's at stage II whereas Eve's overall rate of *na*-constructions is lower at 12.7%.

Table 5.12 presents the distribution of the future particle *tha* (present and omitted) in clauses with verbs.

Table 5.12: Maria's future constructions with particle *tha* used and omitted

| Maria Age | MLU | tha + V | 0tha + V |
|----------------------|------------|--------------------|-------------------|
| 2;0 | 3.8 | 60% (n=2) | 40% (n=3) |
| 2;2 | 3.4 | 75% (n=3) | 25% (n=1) |
| 2;3 | 3.8 | 88.4% (n=23) | 11.6% (n=3) |
| 2;5,4 | 4.2 | 87.8% (n=29) | 12.2% (n=4) |
| 2;5,24 | 3.6 | 91.6% (n=22) | 8.4% (n=2) |
| 2;7 | 3.7 | 85.7% (n=6) | 14.3% (n=1) |
| 2;8,318 | 3.6 | 100% (n=18) | 0% (n=0) |
| 2;8,27 | 4 | 100% (n=13) | 0% (n=0) |
| Total | | 89% (n=116) | 11% (n=14) |

Maria's rate of omission of the future particle *tha* decreases over time. Omission of particle *tha* is at its highest at the ages of 2;0 and 2;2. The numbers, however, are only few. At the ages between 2;3 and 2;7, the rate of omitted future particles fluctuates but with stable decrease. Maria's use of future constructions with the particle *tha* amounts to 15% of all verbal clauses in the data.

Table 5.13 shows the use and omissions of the future particle *tha* in Eve's speech.

Table 5.13: Eve's future constructions with particle *tha* used and omitted

| Eve Age | MLU | tha + V | 0tha + V |
|--------------------|------------|----------------------|-------------------|
| 1;7 | 2 | 100% (n=2) | 0% (n=0) |
| 1;8 | 1.6 | 100% (n=0) | 0% (n=0) |
| 1;9 | 2 | 100% (n=3) | 0% (n=0) |
| 1;10 | 1.9 | 100% (n=10) | 0% (n=0) |
| 1;11 | 2.3 | 100% (n=9) | 0% (n=0) |
| 2;0 | 2.4 | 100% (n=16) | 0% (n=0) |
| 2;2 | 2.5 | 100% (n=23) | 0% (n=0) |
| 2;3 | 2.3 | 100% (n=18) | 0% (n=0) |
| 2;5 | 2.8 | 92.8% (n=13) | 7.2% (n=1) |
| 2;6 | 2.5 | 100% (n=15) | 0% (n=0) |
| 2;9 | 2.8 | 100% (n=9) | 0% (n=0) |
| 2;10 | 2.7 | 100% (n=12) | 0% (n=0) |
| 2;11 | 4.4 | 100% (n=3) | 0% (n=0) |
| Total | | 99.2% (n=133) | 0.8% (n=1) |

Eve's rate of omission of the future particle *tha* is as low as 0.8%. There is only 1/14 missing particle *tha* across all future constructions at the age of 2;5. Eve's rate of use of future clauses is 5.3% out of all verbal constructions in her data.

Examples of future particle *tha* used and omitted are given below:

39. tora tha kathiso edo
now willFUT sit/1sgDEP here

Now, I will sit here

(Maria, 2;5,24)

40. Otha ti (v)gali ti bala
OthaFUT it take-out/3sgDEP the ball

He/she will take-out the ball

(Maria, 2;5,24)

41. tha to majilevo
willFUT it cook/1sgDEP

I will be cooking it

(Eve, 1;10)

42. de Otha fai i Agela
not OthaFUT eat/3sgDEP the Angela

Angela will not eat

(Eve, 2;5)

To summarise so far, modal particles *na/tha* are used in a productive way in both Maria's and Eve's speech. It's difficult to establish if the use of subjunctive and future constructions increases with age because the rate of use is relative to the size of each recording. Omission of the modal particles, however, decreases, with time. The comparison between the use of the particles *na* and *tha* in Maria's and Eve's data

reveals that Maria omits the particle *tha* more than the particle *na*, i.e. 11% vs. 7% respectively, whereas, Eve's omission of the particle *na* is higher than the omission of *tha*, i.e. 1.6% vs. 0.8% respectively.

5.8.3 *RIs*

To establish the use of bare verbal forms in the corpora of Maria and Eve, the following qualitative and quantitative analyses were conducted for each data:

Overuse of the *-i* forms (*-i* form 3sg and other 3sg)

Use of perfective vs. imperfective stem in *-i* forms

Agreement in *-i* forms (non-3sg context)

Null subjects in *-i* forms

5.8.3.1. *Overuse of the -i forms (-i form 3sg and other 3sg)*

VVR have suggested that there is an overuse of the verbal form that involves the suffix *-i*, referred to as the 3rd singular person or participle (VVR, 1996/1998). This form involves predominantly the perfective stem. Examples of 3sg with *-i*⁴⁰ suffix are:

43. ta pets-**i** to matsilari
 thaFUT fall/3sg the pillow
 the pillow will fall down

(Maria 2;0)

⁴⁰ The alternative suffix of 3sg person is *-e*.

44. tha k(l)is-i tin porta
thaFUT close/3sg the door
he/she will close the door

(Maria 2;5,4)

45. tora tha djasvas-i ti kuala
now thaFUT read/3sg the koala
he/she will read the koala (book) now

(Maria 2;5,24)

46. na katharis-i ti miti tu
toSUBJ clean/3sg the nose his
(he/she is going) to blow her/his nose

(Maria 2;7)

47. (th)a (ch)tipis-i
thaFUT hurt/3sg (itself)
it will hurt itself

(Eve 1;10)

48. (th)a me kaps-i
thaFUT me burn/3sg
it will burn me

(Eve 2;0)

49. thelo na chores-**i** mesa
 want/1sg toSUBJ fit/3sg inside
I want (it) to fit inside

(Eve 2;2)

50. theli na kathis-**i** to moraki mu
 want/3sg toSUBJ sit/3sg the babeDIM mine
my little baby wants to sit

(Eve 2;2)

To check VVR's proposed RI stage, an analysis of 3rd singular verbs with different endings (-*i* and -*e*) was conducted in Maria's and Eve's corpora.

Eve's data is transcribed in CHAT, however there is not suffix distinction between -*e* and -*i* for the 3rd singular in the morphological coding. The following CLAN command was used for the identification of 3rd singular forms where the different suffixes were checked manually:

combo +t%mor +s*3*sg* eve.cha

Using this command, all verbs coded 3rd singular were identified in Eve's data. If a 3rd singular form is used in the wrong context then is also coded with an error code. The different verbal endings were checked manually. COMBO was opted for over a FREQ analysis because it provides information about both the verbs and the clause where the context could be checked too.

Repetitions of the same verbal form were counted in the analysis, following the repetition procedure discussed in Methodology. Finally, only clear verbal forms were considered in this analysis i.e. verbs clearly pronounced and identified. Copula *be* was also excluded from this analysis, following the example of almost all studies in early Greek. This is because in Greek (and other languages) copula *be* has both main and auxiliary verb functions (Klairis and Babiniotis, 2004). Copula *be* is used as an auxiliary for periphrastic tenses and with adjectival, nominal, adverbial and prepositional phrases.

In Maria's data the 3rd singular verbal paradigm with different suffixes was checked manually following the above conditions.

Table 5.14 presents the distribution of 3rd singular person verbs divided in verbs ending in *-i* and verbs with different suffix in Maria's speech production.

Table 5.14: Maria's 3rd singular with *-i* and other endings

| Maria | MLU | <i>-i form 3sg</i> | | other 3sg | |
|--------------|-----|-------------------------------------|-----------------------------------|------------------------------------|-----------------------------------|
| | | % <i>-i</i> | % <i>incorrect</i> | % <i>non -i</i> | % <i>incorrect</i> |
| 2;0 | 3.8 | 83.3% (n= 15/18) | 26.6% (n= 4/15) | 16.7% (n= 3/18) | 0% (n= 0/3) |
| 2;2 | 3.4 | 87.5% (n= 7/8) | 14.2% (n= 1/7) | 12.5% (n= 1/8) | 0% (n= 0/1) |
| 2;3 | 3.8 | 76.7% (n= 33/43) | 9% (n= 3/33) | 23.3% (n= 10/43) | 30% (n= 3/10) |
| 2;5,4 | 4.2 | 53.5% (n= 30/56) | 3.3% (n= 1/30) | 46.5% (n= 26/56) | 7.6% (n= 2/26) |
| 2;5,24 | 3.6 | 60.9% (n= 25/41) | 0% (n= 0/25) | 39.1% (n= 16/41) | 18.7% (n= 3/16) |
| 2;7 | 3.7 | 60% (n= 21/35) | 0% (n= 0/21) | 40% (n= 14/35) | 14.2% (n= 2/14) |
| 2;8,3 | 3.6 | 61.5% (n= 8/13) | 25% (n= 2/8) | 38.5% (n= 5/13) | 0% (n= 0/5) |
| 2;8,27 | 4 | 61.1% (n= 33/54) | 0% (n= 0/33) | 38.9% (n= 21/54) | 9.5% (n= 2/21) |
| Total | | 64.1% (n= 172/268) | 6.3% (n= 11/172) | 35.9% (n= 96/268) | 12.5% (n= 12/96) |

In the first three recordings of Maria's data collection (ages 2;0-2;3) 3sg verbs ending in *-i* are used more than 3sg verbs with different suffix, i.e. 76-87% vs. 12-23%. At the age of 2;5,4 the use between verbs with *-i* and other suffix is more balanced, namely 53.5% vs. 46.5%. From 2;5,24 to 2;7 3sg verbs with *-i* are used again more than verbs with other suffix. At the age of 2;8,3 the use of 3sg *-i* form is greater than 3sg with a different suffix, however, there are only few verbs. Finally, at the last recording (age 2;8,27) Maria uses 3sg verbs with *-i* much more than 3sg verbs with a different

ending. Overall 3sg verbs with *-i* are used more than verbs with different suffix, nearly twice as much. Incorrect cases are very limited across Maria's speech production. At 2;0 the rate of incorrect *-i forms* is as high as 26.6% and the gradually reduces overtime, except at 2;8,3 where is as high as 25%. There are only few instances, however, in this recording, and therefore this is not conclusive. For 3sg with a different suffix, incorrect instances are again few in numbers ranging from 0% to the highest pick of 30% at 2;3 and therefore there is no developmental pattern observed. Overall, incorrect cases of 3rd singular verbs are higher in 3sg with a different ending, i.e. 6.3% in verbs with *-i* ending vs. 12.5% in verbs with a different ending. Therefore there is no clear evidence of high rate of incorrect use of 3sg with *-i* suffix.

Table 5.15 presents the distribution of 3rd singular person verbs divided in verbs ending in *-i* and verbs with different suffix in Eve's data collection:

Table 5.15: Eve's 3rd singular with *-i* and other endings

| Eve | MLU | <i>-i form 3sg</i> | | other 3sg | |
|--------------|------------|-------------------------------------|-----------------------------------|-------------------------------------|-----------------------------------|
| Age | | % <i>-i</i> | % <i>incorrect</i> | % <i>non -i</i> | % <i>incorrect</i> |
| 1;7 | 2 | 78.1% (n= 25/32) | 0% (n= 0/25) | 21.9% (n= 7/32) | 0% (n= 0/7) |
| 1;8 | 1.6 | 33.3% (n= 2/6) | 50% (n= 1/2) | 66.7% (n= 4/6) | 50% (n= 2/4) |
| 1;9 | 2 | 68.7% (n= 33/48) | 12.1% (n= 4/33) | 31.3% (n= 15/48) | 6.6% (n= 1/15) |
| 1;10 | 1.9 | 71.4% (n= 60/84) | 10% (n= 6/60) | 28.6% (n= 24/84) | 4.1% (n= 1/24) |
| 1;11 | 2.3 | 72.3% (n= 55/76) | 0% (n= 0/55) | 27.7% (n= 21/76) | 0% (n= 0/21) |
| 2;0 | 2.4 | 86.1% (n= 118/137) | 6.7% (n= 8/118) | 13.9% (n= 19/137) | 0% (n= 0/19) |
| 2;2 | 2.5 | 75.8% (n= 47/62) | 0% (n= 0/47) | 24.2% (n= 15/62) | 0% (n= 0/15) |
| 2;3 | 2.3 | 85.4% (n= 82/96) | 0% (n= 0/82) | 14.6% (n= 14/96) | 0% (n= 0/14) |
| 2;5 | 2.8 | 47.5% (n= 57/120) | 5.2% (n= 3/57) | 52.5% (n= 63/120) | 0% (n= 0/63) |
| 2;6 | 2.5 | 47.4% (n= 28/59) | 3.5% (n= 1/28) | 52.6% (n= 31/59) | 0% (n= 0/31) |
| 2;9 | 2.8 | 70% (n= 7/10) | 0% (n= 0/7) | 30% (n= 3/10) | 0% (n= 0/3) |
| 2;10 | 2.7 | 81.8% (n= 18/22) | 0% (n= 0/18) | 18.2% (n= 4/22) | 0% (n= 0/4) |
| 2;11 | 4.4 | 58.3% (n= 7/12) | 0% (n= 0/7) | 41.7% (n= 5/12) | 0% (n= 0/5) |
| Total | | 70.5% (n= 539/764) | 4.2% (n= 23/539) | 29.5% (n= 225/764) | 4.8% (n= 11/225) |

During the first recording Eve uses 3sg with *-i* more than 3sg with a different suffix. At the age of 1;8 there are only few verbs in 3rd singular. Therefore they are not very informative. From the age of 1;9 to 1;11 the use of 3sg verbs with *-i* is much greater than 3sg with a different ending, i.e. 68-72% vs. 27-31%. The same pattern is observed from 2;0 to 2;3. At the age of 2;5 and 2;6 3sg with a suffix other than *-i* are slightly more used than verbs with *-i* ending, namely 52% vs. 47%. Consequently, from 2;9 to 2;11 3sg with *-i* are again used more than verbs with a different ending. There is however an overall decrease in the use of 3rd singular verbs. On the whole, 3rd singular verbs with *-i* suffix are used much more than 3sg with a different ending, more than half the time.

Incorrect use ranges from 0% to 12.1% for 3sg with *-i* and from 0% to 6.6% for 3sg with a different suffix, across the dataset. At its highest, is 50% at the age of 1;8 for both 3sg with *-i* and 3sg with different suffix. There are only few instances, however, at this recording, and this is not conclusive. Overall, incorrect use of 3sg is the same in verbs ending with *-i* and verbs with a different suffix. Therefore there is no clear evidence of high rate of incorrect use of 3sg with *-i* suffix. Overall, the error rate is very low.

5.8.3.2 *Use of perfective vs. imperfective stem in -i forms*

VVR (1996, 1998) suggested that 3rd singular verbs with *-i* are mostly formed with the perfective stem. To verify this claim a further analysis of the stem distribution across the 3sg with *-i* ending was conducted in Maria's and Eve's corpora.

In MG some verbs do not have morphological stem distinction between imperfective and perfective, so their aspect characteristics are ambiguous⁴¹ (unless used

⁴¹ For a detailed discussion of aspect and single-stem verbs see chapter 4

with one of the particles *na/tha* in which case they are perfectives) and can be checked only from adjacent contextual information. Such information is not often available in children's speech (especially when a verb stands alone in the clause). For this reason those verbs were left out of further analyses. The percentages of these verbs are presented in this analysis for information. Such verbs are referred as *single-stem verbs* (Katis, 1984) e.g. *kano* (= I make), *echo* (= I have), *ksero* (= I know).

Tables 5.16 and 5.17 present the distribution of the various verbal stems across the 3sg *-i* forms in Maria's and Eve's speech production.

Table 5.16: Maria's *-i* forms with aspectual stems

| Maria Age | MLU | perfective | imperfective | (im)perfective (ambiguous) |
|----------------------|------------|-------------------|---------------------|---------------------------------------|
| 2;0 | 3.8 | 6 | 2 | 7 |
| 2;2 | 3.4 | 2 | 1 | 4 |
| 2;3 | 3.8 | 9 | 11 | 13 |
| 2;5,4 | 4.2 | 9 | 10 | 11 |
| 2;5,24 | 3.6 | 7 | 8 | 10 |
| 2;7 | 3.7 | 8 | 5 | 8 |
| 2;8,3 | 3.6 | 2 | 4 | 2 |
| 2;8,27 | 4 | 7 | 11 | 15 |
| Total | | 50 (29%) | 52 (30%) | 70 (41%) |

In Table 5.16 the use of perfective stem with *-i* forms in Maria's speech production is slightly higher than 3sg cases used with imperfective stem i.e. 29%

perfective vs. 30% imperfective. The highest percentage of 3sg is encountered with the use of a single-stem verb i.e. 41%.

Table 5.17: Eve's *-i* forms with aspectual stems

| Eve Age | MLU | perfective | imperfective | (im)perfective (ambiguous) |
|--------------------|------------|-------------------|---------------------|---------------------------------------|
| 1;7 | 2 | 1 | 1 | 23 |
| 1;8 | 1.6 | 1 | 0 | 1 |
| 1;9 | 2 | 5 | 13 | 15 |
| 1;10 | 1.9 | 9 | 24 | 27 |
| 1;11 | 2.3 | 8 | 12 | 35 |
| 2;0 | 2.4 | 35 | 54 | 29 |
| 2;2 | 2.5 | 10 | 23 | 14 |
| 2;3 | 2.3 | 6 | 40 | 36 |
| 2;5 | 2.8 | 21 | 13 | 23 |
| 2;6 | 2.5 | 11 | 8 | 9 |
| 2;9 | 2.8 | 0 | 3 | 4 |
| 2;10 | 2.7 | 2 | 8 | 8 |
| 2;11 | 4.4 | 1 | 2 | 4 |
| Total | | 110 (20%) | 201 (37%) | 228 (43%) |

In Table 5.17 Eve's use of verbs with perfective stem is considerably lower compared to Maria's data i.e. 20% verbs with perfective stem vs. 37% with imperfective. The use of single-stem 3sg verbs is the highest at 43%.

In further summary, there is indeed evidence of overall overuse of the *-i* form 3sg, namely 64% for Maria and 70% for Eve of the total 3rd singular verbs. Regarding aspect, the overuse of *-i* forms is slightly more prominent with the perfective form in Maria's speech. In Eve's speech *-i* forms are used mostly with imperfective stem. Single-stem 3sg verbs are used in a high rate in both Maria's and Eve's corpora. There is not sufficient evidence in Maria's and Eve's corpora of overuse of the perfective stem in 3rd singular verbs with *-i* suffix.

5.8.3.3 *Subject agreement in -i forms (non-3sg context)*

In what follows an analysis of agreement and agreement errors in 3sg *-i* forms with perfective aspect will be performed. Such forms, when produced in the wrong context, are referred to as the Bare Perfective forms (BPs) which are the Greek RIs equivalent.

MG is a null subject language and hence a verb with a phonetically unrealised subject is grammatical. Therefore agreement errors are checked based on the context against the grammaticality of the sentence.

Tables 5.18 and 5.19 present the distribution of BPs in Maria's and Eve's corpora. The columns named *BPs* contain all 3sg *-i* forms with perfective aspect, no or incorrect agreement and no modal particles, whereas the columns named *perfective (other)* contain all other 3sg *-i* forms with perfective aspect that have correct agreement and/or a modal particle.

Table 5.18: Distribution of BPs in Maria's data

| Maria Age | MLU | BPs | Perfective (other) |
|----------------------|------------|---------------|-------------------------------|
| 2;0 | 3.8 | | 6 |
| 2;2 | 3.4 | 1 | 1 |
| 2;3 | 3.8 | 1 | 8 |
| 2;5,4 | 4.2 | 1 | 8 |
| 2;5,24 | 3.6 | | 7 |
| 2;7 | 3.7 | | 8 |
| 2;8,3 | 3.6 | | 2 |
| 2;8,27 | 4 | | 7 |
| Total | | 3 (6%) | 47 (94%) |

Table 5.19: Distribution of BPs in Eve's data

| Eve Age | MLU | BPs | Perfective (other) |
|--------------------|------------|-------------------|-------------------------------|
| 1;7 | 2 | | 1 |
| 1;8 | 1.6 | 1 | |
| 1;9 | 2 | 3 | 2 |
| 1;10 | 1.9 | 4 | 5 |
| 1;11 | 2.3 | | 8 |
| 2;0 | 2.4 | 6 | 29 |
| 2;2 | 2.5 | | 10 |
| 2;3 | 2.3 | | 6 |
| 2;5 | 2.8 | 3 | 18 |
| 2;6 | 2.5 | | 11 |
| 2;9 | 2.8 | | |
| 2;10 | 2.7 | | 2 |
| 2;11 | 4.4 | | 1 |
| Total | | 17 (15.5%) | 93 (84.5%) |

In Tables 5.18 and 5.19, 6% (N=3) in Maria's speech and 15.5% (N=17) in Eve's are *-i* 3sg forms with perfective stem, and no or incorrect agreement and omission of the modal particles (*na/tha*) according to the adult-like model modal construction. These cases are the BP forms. These forms are ungrammatical in the adult language. The rest of the cases all have the particles *na/tha* according to the adult grammar model,

or when no particles are presented (few cases) something else indicates the grammaticality of the sentence (e.g. conditional sentence).

All cases of BPs in both corpora are formed with the perfective stem. Therefore, these BPs are solely formed with the perfective stem as suggested in previous studies. There are, however, only few instances of BPs present in both Maria's and Eve's speech, and hence, the results are not conclusive.

Maria's speech is clearly in a more advanced stage compared to Eve's. The results also indicate that RIs (or BPs in this case) occur for a very short period before the age of 2;5 years. Maria's data suggest that by the age of 2 years Greek children are growing out of any RIs stage they may be. This corresponds between Stage III (MLU: 2 – 2.5) and Stage IV (MLU: 3 – 3.75) in Brown's classification.

Eve's data provide a clearer developmental picture indicating the status of RIs in early Greek. At the age of 1;8 there's only one case of *-i* form used with perfective aspect and incorrect agreement. This is not informative. At the ages of 1;9 and 1;10 3/5 and 4/9 respectively, are cases of *-i* forms with perfective and no agreement (BPs). At the ages of 2;0 and 2;5 there are 6/35 and 3/21 BPs respectively and this designates a reduction of these constructions.

5.8.3.4 *Null subjects in -i forms*

It has been also suggested (VVR, *ibid.*) that BPs in early Greek, i.e. 3sg *-i* form with perfective stem, mostly occur with null subjects contrary to finite verbs, where both null and overt are observed. As discussed in the previous section, MG, being a null subject language, allows verbs with phonetically unrealised subjects.

Based on the previous analysis, the current distribution examines null and realised subjects across 3sg *-i* forms formed with the perfective stem.

In Maria's data all cases of BPs (N=3) involve a null subject as argued in previous studies. As discussed in the previous section, however, Maria's BPs are only 3 cases and therefore the results are not conclusive. In Eve's speech only the 41% of the BPs cases involves a null subject (7/17).

A further analysis on null subjects in the remainder of the verbal paradigm (inclusive of finite forms) reveals that in Maria's data 65% of subjects are null and 35% realised. In Eve's data the rates are similar i.e. 61%. A similar analysis of subjects (both positions for realised subjects) was performed in the adults' speech production in Maria's and Eve's corpora shows that 62% of subjects are null and 38% overt. This result corresponds to the ones found in children's speech.

5.8.4 Other evidence for RIs

More recent studies (Varlokosta, 2002; Hyams, 2001, 2002; among others) proposed some new evidence for the distribution of BPs in early Greek, such as, limited use of past tense, clitic omission and determiner drop in relation to 3sg *-i* forms. In the following section such evidence is examined through a series of analyses of Maria's and Eve's data.

5.8.4.1 Past tense and *-i* forms

Varlokosta (2002) claimed that during Stage 1 (the proposed RI stage), characterised by the overuse of the *-i* forms, tense distinction also tends to be absent (Table 7, Varlokosta 2002: 143). The investigation of past feature across the 3sg forms is relevant in terms of eventivity. Varlokosta (ibid.) suggested that children's use of past tense is limited only to eventive predicates and it is not clear whether past tense indicates

temporal order or the predicates are used with perfective aspect to indicate completion of an event. Hyams (2007), discussed in the first part of this chapter, proposed a very similar premise (Closed Event Hypothesis) using the Linking Algorithm.

In MG event predicates (action/activity) can be modified in time whereas stative predicates cannot (Klairs and Babiniotis, 2004):

O skilos mas gavgizi

Our dog is barking

O skilos mas kimate sto spitaki tou

Our dog sleeps in his dog house

(Klairs and Babiniotis, 2004: 553)

According to Varlokosta (2005), building on Vendler's distinction (1957) between eventive and non-eventive predicates, eventive predicates indicate a dynamic change that takes place within a clearly confined period of time whereas non-eventive predicates refer to situations without a clearly defined beginning or end point. In MG, when a stative verb undergoes *perfectisation* it then acquires an eventive meaning (Giannakidou, 2002).

Based on the above evidence about the presence of past tense, the following analysis investigates the status of past tense in Eve's and Maria's corpora in relation to BPs⁴².

In Maria's data 9.6% of the all the verbs used are in the past tense, whereas in Eve's data the percentage is 13.4%. [+past] morphology in 3sg person context is as high

⁴² For a detailed analysis of past tense in the data collection see chapter 4.

as 41% in Maria's data and 19% in Eve's. Past tense in Maria's and Eve's speech is produced with instance of various verbs, as shown below:

Lithike (= got untied/3sg), *epese* (= fell/3sg), *vrachike* (= got wet/3sg),
irthe (= came/3sg), *epjase* (= held/3sg), *evale* (= put/3sg), *echise* (= poured/3sg),
megalose (= got older/3sg), *potise* (= watered//3sg), *emine* (= stayed/3sg),
djavase (= read/3sg), *petakse* (= threw/3sg), *skupise* (= wiped/3sg),
efaje (= ate/3sg), *teliose* (= finished/3sg), *anevike* (got up/3sg),
jelase (= laughed/3sg).

(Maria – all recordings)

Epesan (= fell/3pl), *vrika* (= found/1sg), ***epeze*** (= **was playing//3sg**),
chtipise (= hit/ /3sg), *erik(s)a* (= dropped/1sg), *va(r)ethika* (= got bored/1sg),
zali(s)tikes (felt dizzy/2sg), *chalase* (= got ruined/3sg), *kolisa* (= got stuck/1sg), *ponese*
(= hurt/3sg), *eftjatse* (= made/3sg), *kopike* (= got cut/3sg),
djavasa (= read/3sg), *efaje* (= ate/3sg), *dografisa* (= drew/1sg),
zestathike (= got hot/3sg), ***ekleje*** (= **was crying/3sg**), *pesame* (= fell/1pl),
idame (= said/1pl), *pigame* (= went/1pl), *ftjatsame* (= made/1pl),
k(l)isame (= closed/1pl), ***malones*** (= **was scolding/2sg**), *zalitikes* (= felt dizzy/2sg),
ek(l)ises (= closed/2sg), *chalases* (= got ruined/2sg).

(Eve – all recordings)

With the exception of few verbs used with imperfective stem (in bold) all other instances of past tense in Maria's and Eve's corpora are eventive predicates formed with the perfective stem. Following Giannakidou (ibid.) the assumption is that in MG

aspect dictates the verbal class, i.e. verbs are unambiguously eventive when they have perfective aspect and stative when they are formed with imperfective aspect.

In Table 5.20 the distribution of verbal stems across past tense verbs is provided for Maria's and Eve's speech production.

Table 5.20: Stem distribution in [+past] verbs in Maria's and Eve's corpora

| | Perfective | Imperfective | Single-stem verbs | Total |
|--------------|-------------------|---------------------|--------------------------|--------------|
| Maria | 95.2% (N=238) | 0.4% (N=1) | 4.4% (N=11) | 250 |
| Eve | 91.5% (N=497) | 3% (N=16) | 5.5% (N=30) | 543 |

The majority of past tense verbs are formed with the perfective stem, i.e. 95% in Maria's and 91% in Eve's speech. The use of imperfective stem is fairly limited, namely, 0.4% and 3% in Maria's and Eve's data respectively. Single-stem verbs were also calculated. Copula was not included in the distribution.

The examples given above suggest that past tense is used from very early in Maria's and Eve's speech. The results in Table 5.20 confirm, however, that past tense is limited to eventive predicates (used with perfective aspect).

51. epese; edo pera ine ta xxx.

Fell/3sg; here are/2pl the/pl xxx

It fell; down here there are the xxx

(Maria, 2;3)

52. ... telika ta pjase i miga ke ta evale epano
...at the end the caught/3sg the fly and the put/3sg up-there

At the end the fly caught them and put them up there

(Maria, 2;5,4)

53. i(rth)se i nona tis
arrived/3sg the godmother her

Her godmother arrived

(Maria, 2;5,24)

54. ochi , (d)e chalase
no, not broke/3sg

No, it didn't brake

(Eve, 1;11)

55. epese i fo(t)ografia
fell/3sg the photograph

The photograph fell off

(Eve, 1;11)

56. chtipise sto magulo

hurt/3sg at-the cheek

She/he was hurt at the cheek

(Eve, 2;2)

5.8.4.2 *Clitic Object Omission in -i forms*

In other languages with RI phenomena a correlation has been observed between the use of non-finite forms and the omission of clitics during this stage (Hamann et al., 1996; Haegeman, 1996). Varlokosta (2002) suggested that such correlation may apply to the Greek equivalent too. Varlokosta (ibid.) reported that 42% of the cases of clitic object omission in obligatory context in Spiros' (1;9) and 83% in Janna's (1;11) involved *-i* forms with no agreement. When *-i* forms with no agreement are compared to other *-i* forms then Spiros' cases of object clitic omission raises up to 57% (Janna's remain the same).

Examples of clitic objects in sentences from Maria's and Eve's corpora are given below:

57. tora t' agapai

now it love/3sg

now (she/he) is loving it

(Maria 2;5,4)

58. mu ta patai i jaja

pronoun/me on them step/3sg the granny

the granny is stepping on them (of mine)

(Maria 2;3)

59. ala den **ti** theli

but not her want/3sg

but she/he doesn't want her

(Maria 2;8,3)

60. **ta** perni to koritsi ke **ta** pai sto saloni

them take/3sg the girl and them bring/3sg to living room

the girl takes them and brings them to the living room

(Maria 2;3)

61. tora, ato tha **ti** (v)gali ti bala

now, this thaFUT her put-out/3sg the ball

and now this will put it out, the ball

(Maria 2;5,24)

62. na **ta** vali i mama

naSUBJ them put-in/3sg the mum

and mum should put them in

(Maria 2;8,27)

63. **to** bazo bano

it put/1sg on-top-of

I put it on top of

(Eve 1;7)

64. kita **to** petao andbaaaa
lookIMP it throw/1sg
look I throw it

(Eve 1;9)

65. **ta** folao edo
them wear/1sg here
I wear them here

(Eve 1;10)

66. ego na **to** ftjatso
I toSUBJ it fix/1sg
I (am going to) fix it

(Eve 2;2)

67. pame na **tis** dume
go/1pl toSUBJ themFEM look/1pl
Let's go to see them (fem)

(Eve 2;5)

68. thelo na **to** chalaso
want/1sg toSUBJ it destroy/1sg

I want to destroy it

(Eve 2;6)

Following Varlokosta's suggestion (ibid.) for the correlation between BPs and clitic omission, Tables 5.21 and 5.22 below present the clitic omission in 3sg and other verbs in Maria's and Eve's corpora respectively.

Table 5.21: Clitic object omission in 3sg vs. other persons in Maria's data

| Maria | | <i>3sg-all</i> | <i>All persons-rest</i> | Total |
|------------------------|---------|-----------------------|--------------------------------|--------------|
| Obligatory contexts | Present | 98 (89%) | 184 (94%) | |
| | Omitted | 12 (11%) | 12 (6%) | |
| Non obligatory context | | 247 | 290 | |
| Total | | 357 | 486 | 843 |

Table 5.22: Clitic object omission in 3sg vs. other persons in Eve's data

| Eve | | <i>3sg-all</i> | <i>All persons-rest</i> | Total |
|------------------------|---------|-----------------------|--------------------------------|--------------|
| Obligatory contexts | Present | 106 (99%) | 306 (96.5%) | |
| | Omitted | 1 (1%) | 11(3.5%) | |
| Non obligatory context | | 917 | 1169 | |
| Total | | 1024 | 1486 | 2510 |

In Tables 5.21 and 5.22 the omission of object clitic in obligatory contexts in 3rd singular verbs in Maria's and Eve's speech production is 11% and 1% respectively

(12/110 and 1/107 respectively). In the remainder of the verbal paradigm Maria's cases of clitic object omission in obligatory contexts is 6% and Eve's is 3.5%.

The proposed correlation focuses, however, on BPs. Tables 5.23 and 5.24 present the omission of clitic object in 3sg *-i* forms without agreement (BPs) and other 3sg *-i* forms (with correct agreement).

Table 5.23: Clitic objects in 3sg *-i* forms (BPs and non) in Maria's data

| | | Perfective Stem | | |
|------------------------|---------|--------------------|------------|--------------|
| Maria | | <i>3sg -i form</i> | <i>BPs</i> | Total |
| Obligatory contexts | Present | 14 (88%) | 2 (67%) | |
| | Omitted | 2 (12%) | 1 (33%) | |
| Non obligatory context | | 34 | 0 | |
| Total | | 50 | 3 | 53 |

Table 5.24: Clitic objects in 3sg *-i* forms (BPs and non) in Eve's data

| | | Perfective Stem | | |
|------------------------|---------|--------------------|------------|--------------|
| Eve | | <i>3sg -i form</i> | <i>BPs</i> | Total |
| Obligatory contexts | Present | 33 (100%) | 0 | |
| | Omitted | 0 (0%) | 0 | |
| Non obligatory context | | 60 | 17 | |
| Total | | 93 | 17 | 110 |

In Maria's speech only 3 cases of 3sg *-i* forms involve omission of object clitic (out of 19) and only one of those is a BPs. In Eve's speech none is present. This is in Maria's speech 33% (one case) of BP involves a clitic omission. There is only one case and this is not very informative. Overall clitic object omission in Maria's and Eve's speech is not very high, contrary to Varlokosta's suggestions for Spiros' and Janna's speech.

Examples of clitic omission in sentences from Maria's and Eve's corpora are given below:

69. 0prt 0cl kutisume pali
 (na/tha) (it) listen/1plDEP again

Let's/we are going to listen to it again

(Maria, 2;0)

70. 0na mi 0cl patisi omos ata
 (na) not (them) step-on/3sgDEP though these

He/she shan't step on them, though

(Maria, 2;3)

71. ochi de tha 0cl sikoso
 no not thaFUT (it/them) lift/1sg

I won't lift it/them up

(Maria, 2;5,4)

72. xxx to ksechasa/1sg ha 0cl v(r)ika

xxx it forgot ha [speech particle] (it/them) found/1sg

I forgot about this, ha! I found it!

(Eve, 1;10)

73. (d)e 0cl chalasa
not (it/them) destroy/1sg

I didn't destroy it/them

(Eve, 1;10)

74. 0prt 0cl taithis
(na/tha) (to) feed/2sg

You should/will feed it

(Eve, 2;2)

5.8.5 *Verb typology and temporal interpretation of BPs*

As mentioned previously, Hoekstra and Hyams (1998) proposed the Eventivity Constraint, according to which RIs are restricted to eventive predicates only. Following Vendler's (1957) verbs' classification based upon time schemata (activity, accomplishment and achievement predicates vs. stative predicates), eventive predicates indicate dynamic change that takes place within a clearly confined period of time whereas non-eventive predicates refer to situations without a clearly defined beginning or end point. Giannakidou (2002) claimed that in MG *perfectivised* stative verbs acquire an eventive meaning.

The notion of telicity also plays a major role in the notion of verb typology. Telic events refer to a change of state that has a result (Smith, 1991) therefore should not be confused with completion. Delidaki (2006) suggested that the perfective

grammatical aspect influences a verb's lexical aspect by assigning the feature of telicity to aspectual classes that do not have it (cf. Giannakidou, 2002); therefore stative verbs become telic when combined with perfective aspect. Activities and accomplishments become telic events when formed with perfective aspect (Delidaki, *ibid.*). The correlation between telicity and perfectivity is also suggested by Christofidou and Stephany (2003) and Stephany and Voeikova (2003) as discussed in the first part of this chapter.

In the *Linking Algorithm* proposed in Hyams (2007) to explain the licensing mechanism of Greek BPs, it is suggested that Greek speaking children produce both telic and atelic BPs, however, these options are not equally available (only 1/4 of BPs are atelic).

Table 5.25 below presents all BPs from Maria's and Eve's analysed data according to their temporal reference (eventivity and telicity).

Table 5.25: Temporal reference of Maria's and Eve's BPs⁴³

| | Eventive | |
|--------------|--|-----------------|
| | Telic | Atelic |
| Maria | vali (= put), patisi (= press), pi (= tell) | |
| Eve | aniksi (= open), besi (= fall), fiji (= leave), feri (= bring), ftjatsi (= make/fix), pi (= tell), vali (= put), kapsi (= burn), fai (= eat), perasi (= cross), pari (= take) | kathisi (= sit) |

In the analysis of Maria's and Eve's BPs in Table 5.23, all verbs are eventive, and all except one is telic, in the sense that it denotes an action that takes place within a

⁴³ Repeated verbs in Maria's and Eve's BPs were omitted from this table.

defined period of time with a specific endpoint. Maria's and Eve's all BPs denote events. In terms of telicity, the prediction that all verbs when combined with perfective aspect become telic has not been borne out from this analysis, since all verbs are eventive and there are no examples of stative verbs (see Gueron, 2008 for a discussion on the differences between perfectivity and telicity).

These findings mostly agree with the results of previous studies mentioned at the first part of this section i.e. Stephany and Voeikova (2003), Varlokosta (2002, 2005). The results also provide support for the Eventivity constraint (Hoekstra and Hyams, 1998) and the predictions of the Linking Algorithm (Hyams, 2007).

5.8.6 *Modal interpretation of BPs*

It has been suggested that RIs with overwhelmingly frequency have modal interpretation (Hoekstra and Hyams, 1998; Hyams, 2002, 2005). Modality expresses necessities, desires and future-like action, and can also convey deontic and volitional meaning (the irrealis sense). This phenomenon in RIs is referred to as the *Modal Reference Effect* (MRE) (Hoekstra and Hyams, *ibid.*).

Stephany's results from her data analyses (1981, 1986) revealed that dynamic verbs are used almost always in modal expressions and mostly with the perfective stem.

Varlokosta (2002) claimed that the majority of *-i* forms involve a future or modal/irrealis interpretation and that this restriction demonstrates the analogy of these forms to RIs in other languages.

Hyams (2002) also considered this analogy to be valid, since like RIs in other languages, and BPs express the child's wish, intention, or need in respect to some eventuality, therefore, has a modal/irrealis interpretation (cf. also Katis 1984, Tsimpli 1992).

No not tell/1sgPERF tell/3sgPERF* down-down

Lit. no, I won't tell; she/he shall/will tell later

[Context: this is Maria's reply to the adult's asking her to tell something to her doll]

Interpretation: modal/future

(Maria, 2;5,4)

78. aniksi

open/3sgPERF*

Lit. She/he shall/will open (it)

[Context: the adult asked Eve if she wants to open something warning her that her mother might get upset with her if she does; utterance before opening of object]

Interpretation: modal/future

(Eve, 1;8)

79. ochi (m)i besi, kalo

no not fall/3sgPERF*, good

Lit. this should not fall; it's good

[Context: Eve is referring to the tape recorder after this has fallen]

Interpretation: past

(Eve, 1;9)

80. (m)i bes(i) ochi bubum

Not fall/3sgPERF no bumbum (referring to the noise of fallen object)

Lit. this should not fall

[Context: Eve is referring to the tape recorder after this has fallen]

Interpretation: past

(Eve, 1;9)

81. i Eva fiji ena xxx

the Eve leave/3sgPERF* one xxx

Lit. Eve will/shall leave, one xxx

[Context: after this utterance Eve is asking for her shoes to go out for a walk]

Interpretation: modal/future

(Eve, 1;9)

82. kathisi

sit/3sgPERF*

Lit. She/he will/shall sit

[Context: Eve's mother is asking if she can sit next to Eve; utterance before her mother's sitting down]

Interpretation: modal/future

(Eve, 1;10)

83. kathisi i Eva , xxx edo

sit/3sgPERF* the Eve, xxx here

Lit. Eve will/shall sit here

[Context: the adult lays a blanket on the floor; utterance before Eve sits on the blanket]

Interpretation: modal/future

(Eve, 1;10)

84. feri (so)kolata ti Evis

bring/3sgPERF chocolate-bar for Eve

Lit. she/he shall/will bring a chocolate bar for Eve

[Context: Eve explains that the man in the book goes to work and he brings a chocolate bar for her when he returns]

Interpretation: modal/future

(Eve, 1;10)

85. fiji

leave/3sgPERF*

Lit. she/he will/shall leave

[Context: this is Eve's reply to the adult asking her to sit down; after the utterance, when asked where she wants to go Eve says that she wants to go out to the playground]

Interpretation: modal/future

(Eve, 1;10)

86. e(g)o ftjatsi xxx to

I make/3sgPERF* xxx the

Lit. I shall/will make xxx the

[Context: the adult asks Eve to make a toy-house; Eve replies this utterance before the toy-house is constructed]

Interpretation: modal/future

(Eve, 2;0)

87. ena kathisi eki mesa

one sit/3sgPERF there inside

Lit. she/he will/shall sit inside there

[Context: Eve is playing with a toy-house and little people; utterance before the sitting down to the toy in the house]

Interpretation: modal/future

(Eve, 2;0)

88. to Thoma pi

to Thomas tell/3sgPERF*

Lit. I shall/will tell Thomas

[Context: Eve is aware that the recording is intended for me and often speaks directly to me; Eve wants to tell me about her parent planned holidays]

Interpretation: modal/future

(Eve, 2;0)

89. va(1)i (e)do mesa

put/3sgPERF* here inside

Lit. She/he shall/will put it inside here

[Context: Eve is about to put back in a container a candy that she had in her hand and became sticky after the adult's suggestion; utterance before the actual action]

Interpretation: modal/future

(Eve, 2;0)

90. kathisi ato eki

sit/3sgPERF this there

Lit. this shall/will sit there

[Context: Eve is forcing her doll to sit on a toy-chair where it does not fit; Eve replies to the adult asking what is she trying to do; utterance before successful sitting]

Interpretation: modal/future

(Eve, 2;0)

91. ke me kapsi aaa
and me burn/3sgPERF uh-oh

Lit. and this shall/will burn me

[Context: Eve is referring to a toy red hot chilli pepper she has in her toy shopping basket; utterance before the action of burning]

Interpretation: modal/future

(Eve, 2;0)

92. de fai i Agela
no eat/3sgPERF the Angela

Lit. Angela will/shall not eat

[Context: Eve's reply to the adult asking what did Angela had to eat]

Interpretation: past

(Eve, 2;5)

93. na mu to chadepthi ke mu perasi
to me it pet/3sg and me pass/3sgPERF

Lit. she/he shall pet it for me and it will pass

[Context: Eve is referring to an accident she had in the past when she hurt herself]

Interpretation: past

(Eve, 2;5)

94. mi mu to pari

no me it take/3sgPERF

Lit. she/he will/shall not take it from me

[Context: the adult is looking at Eve's book; utterance before the actual action]

Interpretation: modal/future

(Eve, 2;5)

In the examples (75) – (94) above the context of Maria's and Eve's BPs utterances was checked. Without going into great details about the exact meaning of each utterance i.e. wish, desire etc., the context analysis reveals that 15/20 of all BPs in the corpora appear to have modal/future interpretation; four of them (examples 79, 80, 92, and 93) refer to past events and one (example 75) is unclear but with a possible modal/future connotation. The context analysis partly supports the claim made by most studies previously discussed, that Greek children's BPs predominantly express irrealis reading by conveying a modal/future interpretation. BPs, like RIs in other languages, are mostly restricted in what regards their temporal reference, to modal interpretation.

5.8.7 *A note on usage-based analyses*

Usage-based studies on RIs have mostly provided results of tailor-made computer based programmes⁴⁴ that simulate children's speech, e.g. MOSAIC. Such programmes are not publicly available and often their methodologies are designed around their study, thus offering, limited ground for replicating the analyses.

It was not possible to use the MOSAIC model in this study. The following discussion, however, sheds some light on usage-based accounts from a developmental point of view.

MOSAIC, as discussed at the first part of this chapter, is used with the aim to simulate children's speech, the prediction being that RI phenomena observed in various languages, can be the result of children omitting sentence-internal elements (Freudenthal, et al., 2005). This is the direct result of frequency effects, i.e. children do not process low frequency items, usually sentence internally; therefore input drives the learning mechanism. The assumption is that the difference of the RI status between languages is due simply to surface characteristics of a language.

The methodology proposed in Freudenthal et al., using MOSAIC cannot be replicated, since Greek is a pro-drop language with relatively free word-order. Therefore, the predictions for Greek are slightly different: RIs are reduced forms (omission of sentence internal part) of the subjunctive in adult MG. Let us consider the following example of RI from Eve 1;8, repeated below:

95. aniksi

open/3sgPERF*

Lit. She/he shall/will open (it)

⁴⁴ Similar programmes e.g. Autotracer, not publicly available, are used for other usage-based analyses such as trace-back.

According to MOSAIC simulation of children’s speech analysis, Eve’s utterance in (95) should be the result of omission of sentence internal parts of the adults’ subjunctive utterances. The following examples (96 and 97) illustrate adult structures of subjunctive in main and embedded sentences from the same recording with the same verb produced before Eve’s utterance in (95):

96. (ela) n’ aniksume afto
 (Come) to open/1plPERF this

Let’s open this – shall we open this!

(Eve, 1;8)

97. thes n’ aniksis eki-pera?
 want/2sg to open/2sgPERF there?

Do you want to open this there?

(Eve, 1;8)

When the subject is not phonetically realised, the verb of the subjunctive construction in (96) is in the middle of the utterance (being a main sentence) following the modal particle *na*. Similarly, when the subjunctive is embedded as in (97) the subjunctive verb follows the main verb, in the middle of the utterance. From (96) and (97) it can be said that verbs in Greek should be more prominent than in non-pro-drop languages, like English. Also, auxiliaries are less frequent in Greek than in English, since tense is marked in the main verb. Therefore, there is less scope for omitting words in the middle of the sentence.

Children's RIs do not appear to be the result of reduced adult input i.e. an utterance missing the initial and the final part. All adult sentences used with the same verb were checked and the results show that adults have not used the exact form, i.e. *aniksi* in the same recording (Eve, 1;8). The examples presented above actually provide counter evidence for the prediction of MOSAIC.

Kupisch and Rinke's (2007) cross-linguistic study on RIs confirmed the suggestions that RIs in children's speech is not the result of token frequency of infinitives (adult like or not) in the input. Their study of adult and children's data in languages such as Italian, Brazilian Portuguese, French, German and English (discussed at the beginning of the chapter) suggested that adult input does not make children produce RI (Kupisch and Rinke, *ibid.*). Their results rather suggested that 'syncretism and ambiguous verb morphology cause delays in the production of finite verb, i.e. Tense morphology' (Kupisch and Rinke, 2007:105); thus the morphological properties of a language may be related to the rate of RIs produced by children learning this language (cf. Legate and Yang (2007) and Blom and Wijnen (2006), according to which the RI-phenomenon in child language can be considered to be a morphological learning problem related to ambiguities in the input).

Kupisch and Rinke's proposal (2007) may provide some kind of resolution when linked to some morphological ambiguities on MG discussed at the first part of this chapter and summarised here.

MG has a rich verbal system where tense morphology is expressed in many different ways. There is, however, an ambiguous form used in two distinct constructions with a different morphological role and different meaning, explained below:

- i. The past participle with the perfective stem used for the formation of periphrastic tenses such as perfect, pluperfect and future anterior with the employment of an auxiliary verb which is inflected (see section 4.3 and section 5.4 for examples). Perfect is a telic tense and as such expresses telicity in the past (an action with end point in the past) whereas the Pluperfect indicates that one action happened before another action took place and the latter action is providing a point of reference (Klairis and Babiniotis, 2004).
- ii. 3rd singular aorist subjunctive form with perfective aspect used with the modal particles *na/tha* preceding the verb (see section 5.4.2 for examples). The subjunctive expresses wish, desire, request, order, permission etc. from the speaker's point of view (Holton et al., 1997).

This form has been suggested to be the RI analogue in child Greek (cf. section 5.5.2). Based on Kupisch and Rinke's (ibid.) proposal it can be suggested that Greek speaking children's production of non-adult constructions is the result of this morphological ambiguity. The fact that MG has a rich morphological system for the expression of tense does not cause a learning problem, but it may, rather facilitate the acquisition process and may also explain why the RI stage in child Greek appears to be so short. This single ambiguous form, however, may cause a brief learning problem to children and it may be considered to cause RIs production.

5.9 Summary and discussion

This chapter's aim was to examine the use of finite and non-finite verbal forms in early Greek through the analysis of data from Maria's and Eve's speech productions, focusing on the phenomenon of RIs in children's early speech. The analyses performed in this

chapter intended to answer two research questions: firstly, whether children use non-finite verbal forms in main contexts in their early speech, that later disappear. From a usage-based point of view, a second question is how these early non-adult constructions are related to the adult input, i.e. are they the result of some adult structures.

The predictions for the first question are that children's use of non-finite non-adult structures develops in two stages; during stage I these forms are overused whereas during stage II such forms are used less and they eventually disappear from children's speech. The predictions to the second question are dual; either the adult input has a frequency effect to children's speech in the sense that children use some kind of reduced sentence produced by the adults or that morphological ambiguities in a given language may be the cause of the production of such sentences in children's speech.

A variety of morphological and syntactic features were analysed based on previous studies e.g. Stephany, Varlokosta, Hyams, to mention but a few, as summarised below.

An error analysis was completed to check the errors performed in Maria's and Eve's speech that might be related to non-finite forms, e.g. missing modal particles. The error rate in Maria's and Eve's data was very low and it appears to be at similar levels reported for other children in other languages.

A modal particles analysis was carried out in order to establish the rate of use of subjunctive and future constructions and the rate of omission of such particles in obligatory contexts. The findings of the analysis suggest that the particles *na/tha* are used very early in Maria's and Eve's speech. Modal particles are often omitted in obligatory contexts, however, this pattern decreases over time.

It has been suggested (VVR, 1998) that there is a high use of incorrect 3sg verbal forms with *-i* suffix mostly used with perfective aspect by Greek speaking

children. These forms are referred to as BPs. An analysis of overuse of the *-i* forms (*-i* form 3sg and other 3sg) was performed to test the existence of two stages where the overuse of *-i* forms decreases with time. The analysis provides evidence of overuse of the *-i* form 3sg.

A stem analysis was performed to check if *-i* forms are used more with the perfective stem rather than the imperfective one. The analysis suggests that such overuse is slightly more prominent with the perfective form only in Maria's data. In Eve's speech *-i* forms are used mostly with imperfective stem. Single-stem 3sg verbs are used in a high rate in both Maria's and Eve's corpora.

An analysis of agreement of *-i* form was conducted in order to identify those *-i* forms used in non-3sg context. The analysis suggests that only few instances, i.e. 6% in Maria's speech and 15.5% in Eve's, are *-i* 3sg forms that have no or incorrect agreement.

An analysis of null subjects in *-i* forms was carried out to establish whether *-i* forms occur with null or realised subject. In Maria's data all cases of BPs (N=3) involve a null subject whereas in Eve's speech only 41% of the BPs cases involve a null subject (7/17).

In terms of stages for the presence of such forms in Maria's and Eve's speech, all the analyses above indicate that BPs occur for a very short period before the age of 2 years; after the age of 2;5 these constructions are fading away. This corresponds between Stage III (MLU: 2.50 – 2.99) and Stage IV (MLU: 3.0 – 3.99) in Brown's classification.

To summarise so far, Maria and Eve use a bare non-finite verbal form, with perfective aspect. This form is ungrammatical in the adult language. There are only few instances of these bare forms i.e. 3 in Maria's and 17 in Eve's data and therefore this

'stage' can be considered very brief. Both Maria and Eve are growing out of this use by the age of 2;5 years.

Further evidence for the existence of a RI stage in early Greek proposed in previous studies (cf. Varlokosta, 2002) was also analysed. Analyses of Past tense and Clitic Object Omission in relation to *-i* forms was completed to test previous empirical data. The suggestions that past tense is not productively used during this stage was not borne out from the analysis of Maria's and Eve's data. The results suggest that past tense is productive from very early in language acquisition and tense distinction is in operation, although limited to eventive predicates. The suggestion that clitic omission is high during this stage was also not borne out from the analysis of Maria's and Eve's data.

In terms of temporal interpretation, Maria's and Eve's BPs are restricted to eventive predicates (except one case) and all denote telicity. A context analyses also revealed that most BPs appear to have modal/future interpretation expressing irrealis; some BPs however, refer to past events. This partly provides support for the Eventivity Constrain (Hoekstra and Hyams, 1998).

From a usage-based point of view, an examination of BPs in Maria's and Eve's data suggests that the use of those forms does not seem to be the result of reduced adult input. The data actually provide counter-examples for the prediction of MOSAIC. Based on Kupisch and Rinke's (2007) study the fact that Greek is a rich morphologically language providing various ways to express tense, may explain the speed of acquisition. At the same time, the presence of an ambiguous form in this system may be the cause of the brief use of non-adult form during the acquisition process.

Taking into account all the above, it appears that Maria's and Eve's speech is at a very advanced stage of development. There seems to be a short stage during which Maria and Eve employ the use of non adult forms in order to express modality or past. It is difficult to identify which constructions of the adult language these forms correspond to (participles or subjunctive) and exactly what these forms try to express. However, the ambiguity between two forms and two meanings may cause some kind of confusion to children trying to figure out the exact mapping of this form. Such a non-target use is very brief and children, very swiftly grow out of it.

6 PERSON AND NUMBER MARKING WITHIN THE VERBAL DOMAIN

6.1 Aims of the chapter

This chapter deals with the development of person and number verb morphology and subject-verb agreement. This chapter will consider previous theoretical and empirical data about the acquisition of subject-verb agreement and thereafter, will examine the acquisition of person and number morphology in Maria's and Eve's data. The aim of this chapter is to examine these inflectional features in Maria's and Eve's speech production by addressing the following questions:

1. When does the distinction between 1st, 2nd and 3rd person emerge (first use)?
2. When is the distinction between 1st, 2nd and 3rd person acquired?
3. When does the distinction between singular and plural [number] emerge (first use)?
4. When is the distinction between singular and plural [number] acquired?
5. Do children substitute 1st and 2nd persons with 3rd?
6. Do children substitute plural with singular?
7. Do children use person/number marking productively from the onset? If not, when do they use person/number marking productively?
8. Are more frequently used verb forms of person/number in adults more frequent in children too?

Based on the literature review and acquisition studies considered in the first part of this chapter, the predictions are as follows:

- For questions 1 and 2 the prediction is that 3rd emerges/is acquired before 1st and 2nd; for questions 3 and 4 singular emerges/is acquired before plural.
- For questions 5 and 6 the prediction is that children substitute 1st and 2nd person with 3rd and plural with singular.
- For question 7 the prediction is that children do not use person and number marking productively from the onset.
- For question 8 the prediction is that children use more frequently verb forms used frequently in the adult language they hear.

In terms of theories, the generative approach predicts for questions 1, 2, 3, and 4 (ordering of person and number markings) that children acquire inflectional morphology, namely the features of [person] and [number] very early (cf. VEPS, VEKI). For the questions 5, 6 the generative model predicts that children do not substitute 1st and 2nd person with 3rd and plural with singular. Weaker accounts of maturation, however (e.g. Tsimpli, Katis), suggest that children at an early stage substitute 1st and 2nd with 3rd person and plural with singular. For question 7 the generative model predicts that children use the inflectional paradigm productively (both features). Finally, for question 8, the generative model makes no predictions.

Usage-based theories make the following predictions: For questions 1, 2, 3, and 4 children's speech is defined by three stages, i.e. the pre-, the proto and the morphology proper stage. Therefore, acquisition is not instant, but subject to a gradual development. For questions 5 and 6 usage-based theories predict that frequently used verbs in the adult language will be frequently used in children's speech too; therefore, no substitution errors are predicted. For question 7 the prediction is that children do not use person and number morphology productively from the onset (cf. Verb Island

Hypothesis). For question 8 usage-based theories predict that children's frequency of use of verbs is closely related to the frequency adults use these verbs (cf. network model).

Xanthos' et al. hypothesis about the effect of a language's rich morphology to the speed of acquisition makes no predictions for the questions mentioned above; their hypothesis, however, will be discussed at the end of the chapter.

The first part of the analyses' section presents quantitative data of each child in terms of accuracy of person and number marking of verbs including examples. The second part presents error analyses including examples of incorrect use of verbal morphology. Finally, the last section contains qualitative analyses and discusses the productivity of person and number marking and subject-verb agreement; also, from a usage-based point of view, frequency analyses are presented and analyses in relation to the Verb Island Hypothesis.

6.2 Introduction

The acquisition of verbal agreement has been previously addressed by exploring if verbal agreement morphemes, i.e. person and number, are used in appropriate contexts and consistently in children's speech.

It has been shown that plural morphology is initially absent in language development and hence plural agreement morphemes appear some months after singular ones (Guasti, 2002; Hyams, 1986; Pizzuto and Caselli, 1992). Guasti (2002) showed that plural morphology appears some months after the emergence of singular one. According to Guasti (*ibid.*), this phenomenon is not limited to plural verbal inflection, but seems to reflect a more general delay in the use of plurality. In early Italian at around age 2, the use of singular morphemes amounts to 90% correct in obligatory

contexts (Pizzuto and Caselli, 1992). Since contexts for the use of plural inflections are initially limited, plural person markers appear later (by some months) than singular ones.

It is also very often argued in the literature that agreement errors are rare and mostly occur with plural subjects. This has been observed in a variety of languages; in early Italian (ages range from 1;4 to 3;0) agreement errors are rare, about 3-4% (Guasti, 1994, Pizzuto and Caselli 1992). Torrens (1995) reported an error rate of about 1.72% in early Catalan and Spanish. In early German (Andreas; age 2;1), Poeppel and Wexler (1993) found that at around age 2 children use the first and third singular agreement morphemes accurately; the second singular morpheme and the three plural morphemes are rare or absent in his speech but agreement errors are rare. Also in early German Clahsen and Penke (1992), in their study of the Simone corpus considered the 2nd singular suffix *-st* to be the milestone for the acquisition of person and they position its productive use only after 2;4,17. Meisel and Ezeizabarrena (1996), in their study on agreement in early Basque (age range of three children from 1;9 to 4;0) found that subject as well as object agreement markings are used almost without errors from the very beginning, as far as the choice of the correct person features is concerned. They have identified some problem areas with plural marking at the beginning but this does not necessarily cause the children to omit affixes or use alternative person marking.

These results are compatible with the claim that correct agreement features on verbal inflectional morphology are acquired very early. To account for these data Hoekstra and Hyams (1995) proposed the hypothesis of Early Morphosyntactic Convergence (EMC) and in a similar way Wexler (1998) proposed the hypothesis of Very Early Knowledge of Inflection (VEKI).

Table 6.1 presents the rate of agreement errors in finite contexts in a variety of languages:

Table 6.1: Rate of Agreement error in children's languages

| Child | Language | Age | N | %error | Source |
|-----------|-----------|----------|------|--------|--------------------------|
| Simone | German | 1;7-2;8 | 1732 | 1% | Clahsen and Penke 1992 |
| Martina | Italian | 1;8-2;7 | 478 | 1.6% | Guasti 1994 |
| Diana | Italian | 1;10-2;6 | 610 | 1.5% | Guasti 1994 |
| Guglielmo | Italian | 2;2-2;7 | 201 | 3.3% | Guasti 1994 |
| Claudia | Italian | 1;4-2;4 | 1410 | 3% | Pizzuto and Caselli 1992 |
| Francesco | Italian | 1;5-2;10 | 1264 | 2% | Pizzuto and Caselli 1992 |
| Marco | Italian | 1;5-3;0 | 415 | 4% | Pizzuto and Caselli 1992 |
| Marti | Cat/Spain | 1;9-2;5 | 178 | 0.56% | Torrens 1995 |
| Josep | Cat/Spain | 1;9-2;6 | 136 | 3% | Torrens 1995 |
| Gisela | Catalan | 1;10-2;6 | 81 | 1.2% | Torrens 1995 |
| Guillem | Catalan | 1;9-2;6 | 129 | 2.3% | Torrens 1995 |

(Adapted from Hoekstra and Hyams, 1998, p. 84)

Table 6.1 (repeated here for convenience) shows that the rate of agreement error in a variety of children's languages is 4% or lower. Low rates of agreement error support further the idea that children have knowledge of agreement features very early.

The present study investigates the status of these features in the initial grammar of Greek, through a series of analyses of two new data bases of spontaneous speech. The aim of the analyses is to identify if the Inflectional Phrase (IP) in children's language is specified for these features.

6.3 Literature Review

In chapter 2 the two main models of acquisition were discussed in detail. The following sections present what these two models suggest in terms of inflection. The first part starts with the generative approach suggesting that inflection is acquired very early and is subject to the setting of parameters (Wexler, 1996, 1998; Hoekstra and Hyams, 1995, 1998; Hyams, 1996, 2008), while the second part discusses the usage-based model which rejects the existence of rules and explains the process of acquisition of inflection based on input, frequencies and their connections (Tomasello, 2003; Bittner et al., 2003, Bybee, 2007, Xanthos et al., 2009).

6.3.1 Very Early Parameter Setting and Very Early Knowledge of Inflection

Wexler (1996) proposed that, at around 18 months of age, children set correctly the basic parameters. This is known as the Very Early Parameter Setting (VEPS). Wexler (ibid.) argued that certain aspects of morphosyntactic inflectional development emerge later in children's development, although parameters are set very early in language acquisition.

Wexler (1998) also proposed the Very Early Knowledge of Inflection (VEKI) according to which children know the grammatical and phonological properties of inflection components in the language they learn at the earliest observable stage (from the time that the child enters the two-word stage around 18 months of age). It has also been suggested that agreement morphemes are known to very young children (Wexler, 1992, 1994).

Poeppl and Wexler (1993) argued that young German-speaking children do not make agreement mistakes; if they produce a verbal inflection then the subject appropriately agrees with it. Equally, if a subject carries certain features, then the

inflection on the verb will carry the same features (Poeppel and Wexler, 1993). Poeppel and Wexler (ibid.) showed that the number of 3sg verbs with first person subject is very rare. For English, Harris and Wexler (1996) showed that *-s* occurs much more frequently when the subject is third singular than with other persons.

6.3.1.1 The Stem Parameter (Hyams, 2008)

For the acquisition of inflectional morphemes, Hyams (2008) proposed the Stem Parameter, according to which a stem is/is not a well-formed word in the language. Children, Hyams (ibid.) claims, set such parameter early and the setting of it affects the time and manner of the acquisition of inflection morphology in the language they learn.

Hyams (ibid.) proposed that the inflectional component of a given language is part of either the core or the periphery of its grammar; core inflectional components are easier to acquire than peripheral ones. Core inflectional systems are closely determined by UG parameters whereas peripheral inflectional systems are subject to exceptions, deviations and relaxations of parameter setting (Hyams, ibid.). The Stem Parameter is set by the child based on experience, i.e. the input.

As a result of this, Hyams (ibid.) claimed⁴⁵ that children learning a language rich in inflection, acquire its inflectional system early and with fewer errors than children learning a language poorer in inflection.

6.3.2 Usage-based models

The central principle of usage-based models of language acquisition is that language structure emerges from language use (Tomasello, 2006). This theory of language acquisition is a single-process system where children acquire more regular constructions

⁴⁵ Hyam's view is reflected in Xanthos' et al. study, discussed later in this chapter.

in the same way they acquire more complex constructions, simply by learning them (Tomasello, 2006).

According to Tomasello (*ibid.*), children initially learn concrete pieces of language, of many different shapes and sizes; later on children generalise and construct more abstract linguistic constructions – which underlie their ability to generate creative new utterances.

Within this framework lie the pre-, proto- and morphology proper stages of acquisition of inflection, discussed in the following section. Also discussed in the following section is the Verb Island Hypothesis (Tomasello, 1992) and the notion of morphological richness from Xanthos' et al. (2009) study, both developed from a usage-based point of view.

6.3.2.1 *Pre-, Proto- and Morphology proper*

Usage-based approaches of language acquisition proposed that children's morphological development is divided into three stages⁴⁶, namely pre-morphology, proto-morphology, and morphology proper (Dressler and Karpf, 1995; Dressler, 2003; Bittner et al., 2003 amongst others).

During the pre-morphological stage, children are thought to learn language by rote while they only start constructing morphological patterns by analogy at the proto-morphological stage. Adult like morphology is only achieved at the morphology proper stage where children create complex morphological constructions (Dressler, 2003; Bittner et al., 2003).

Along the same lines with other usage-based approaches, the above mentioned morphological stages are based on the assumption that children select highly frequent

⁴⁶ A more detailed discussion about the pre- proto- and morphology proper stages can be found on chapter 2.

items from what they hear (principle of pattern selection) and then construct patterns as a result of this natural selection from the input (Bittner et al., 2003).

6.3.2.2 *Verb Island Hypothesis*

The Verb Island Hypothesis (cf. chapter 2, for a more detailed discussion) claimed that children learn the combinatorial possibilities and the syntactic marking for each verb individually (Tomasello, 1992). In this sense, children's early syntactic competence is local and comprises of independent verb island constructions with few or no structural relationships between them (Tomasello, 1992, 2006).

Children are able to generalise patterns and transfer their knowledge from one verb to another only later in their development, i.e. create categories of verbs in which structural characteristics can be applied across groups (Tomasello, *ibid.*).

6.3.2.3 *Child-directed speech, inflectional development, and morphological richness*

Xanthos et al. (2009) in their study on morphological richness in the early development of nominal and verb inflection investigated the relation between the morphological richness of child-directed speech and its impact in the development of inflection in children's speech.

Xanthos' et al. (*ibid.*) study introduced a new methodology to investigate the question above comparing the rate of inflectional development in children acquiring languages of different morphological typology and with different degrees of morphological richness.

According to previous studies (Comrie, 1981; Greenberg, 1954, Dressler, 2005 etc.), the morphological richness of a given language's inflection system is considered

to comprise of two components: the syntagmatic and the paradigmatic one. Syntagmatic richness is when a number of inflectional affixes can be combined in a single word (Comrie, 1981; Greenberg, 1954). For instance, English verbs can be combined only with one marker e.g. *play-ed, play-s*, whereas Greek verbs can be combined with tense, person, and number markers e.g. *e-graps-an* (= wrote/3pl). Thus, Greek is syntagmatically richer than English. Paradigmatic richness is when a given language has a large number of inflected forms per lemma (Dressler, 2005). For instance, English nouns can be inflected only for number e.g. *table – tables* while Greek nouns have four different inflected case forms in the singular and three in the plural, as shown below:

Singular

| | |
|----------|---------------------------------------|
| o likos | wolf +masculine +nominative +singular |
| tu liku | wolf +masculine +genitive +singular |
| ton liko | wolf +masculine +accusative +singular |
| Like | wolf +masculine +vocative +singular |

Plural

| | |
|-----------|-------------------------------------|
| i liki | wolf +masculine +nominative +plural |
| ton likon | wolf +masculine +genitive +plural |
| tus likus | wolf +masculine +accusative +plural |
| Liki | wolf +masculine +vocative +plural |

Thus, English is paradigmatically less rich than Greek.

Building on the morphological richness components mentioned above, Xanthos and Gillis (2010) proposed an alternative way of defining morphological richness in

terms of the average number of inflected word-forms per lemma (Xanthos et al., 2009) namely the *Mean Size of Paradigm* (MSP) as follows:

MSP = F/L , where F stands for the number of distinct inflected word-forms in a sample and L for the number of distinct lemmas. Thus, given the sample 'has, are, have, has, are', containing 5 inflected English verb forms (tokens), one finds $|L| = 2$ (HAVE and BE), and $|F| = 3$ (has, have and are), so that $MSP = 3/2 = 1.5$.

(Xanthos et al., 2009, p. 9)

Xanthos' et al. (ibid.) study analysed data from children between 1;3-3;0 years old, acquiring a range of nine typologically diverse languages all using suffixes in inflecting nouns and verbs. The languages are divided according to their morphological richness as follows:

- Weakly inflecting languages: French, Dutch and German
- Strongly inflecting languages: Russian, Croatian and Greek
- Agglutinating languages: Turkish, Finnish and Yucatec Maya⁴⁷

(Adapted from Xanthos et al., 2009, p. 10)

The results of the study showed that children exposed to morphologically rich languages acquire morphology at a greater speed compared to children learning morphologically less rich languages: Xanthos et al. (2009) argued that variation in child-directed speech has a positive effect on children's early development of inflection,

⁴⁷ A language of Mexico

i.e. child-directed speech of morphologically rich languages plays a major part in the speed of morphological development in children's speech. Exposure to a morphologically rich language might appear to create complications to the acquisition of the language because of the complexity of the language; however, according to Xanthos et al. (ibid.), children exposed to input rich in inflection learn different forms and meaning quicker than children exposed to a poorer inflectional input.

6.4 Person and number in Modern Greek

Modern Greek is rich in terms of both inflectional and case morphology. Agreement specification includes both person and number marking in combination with tense, mood, and aspect marking.

Greek verbs have inflectional morphology that marks three persons and two numbers (Klairis and Babiniotis 2004; Holton et al., 1997). Below are the inflectional suffixes of the agreement paradigm for the two conjugations in the Active Voice:

| Person | 1 st conjugation | | 2 nd conjugation | |
|-----------------|-----------------------------|----------|-----------------------------|---------------------|
| | Singular | Plural | Singular | Plural |
| 1 st | -o | -ome/ume | -(a)o, -o | -ame/ume, - ume |
| 2 nd | -is | -ete | -as, -is | -ate, -ite |
| 3 rd | -i | -une | -a(i), -i | -an(e)/un(e), un(e) |

According to Holton et al. (1997), there are three persons i.e. 1st, 2nd and 3rd and two numbers i.e. singular and plural. The person and the number of the verb agree with the person and the number of the subject noun phrase (Holton et al., ibid.):

Ego taiz-o tis kotes to proi
I-1sg feed/1sg the chickens in the morning

I feed the chickens in the morning

Aftos potiz-i ton kipo kathe mesimeri
He-3sg waters/3sg the garden every noon

He waters the garden every noon

6.5 Acquisition of verbal inflection in MG

6.5.1 Number and Person Markers (Katis)

Katis (1984), in her study on the acquisition of the Greek verb, used one longitudinal sample of a Greek speaking girl, Marilena, between the ages of 2;6 – 4;0, alongside cross-sectional data from 21 children of an age range of 2;0 to 4;11.

From the cross-sectional sample, Katis (ibid.) observed that in reference to number markers, plural is not provided in some of the 2-year old children and one child aged 2;7; amongst them 3pl marker is the main plural marker one used in children's speech. Katis (ibid.) also observed, from the cross-sectional sample, that in children's utterances, in plural errors, the subject is always marked as plural but not the verb. Katis suggested that the plural morpheme is phonemically marked compared to the unmarked singular since it requires an extra syllable to be added.

The patterns observed in person and number markings suggest that singular emerges earlier and is more frequent than plural, and Katis proposed the following development stages:

- i. Attestement of the basic 3sg form only
- ii. Differentiation between 1sg and 2sg
- iii. Differentiation between the 2sg and 3sg as well

(Katis, 1984, p. 98)

According to Katis (ibid.), the order of emergence of the person and number categories, in her analysed data, is singular before plural and 3rd person before 1st and 2nd.

At the age of 2;7, one of the children from the cross-sectional sample shows a preference in using the unmarked 3sg but some 1sg forms are also provided (Katis, ibid.). Katis (ibid.) also observed that children do not demonstrate any problems understanding the contrast between 1sg and 3sg, but some of them do not always produce 2sg, e.g. one child at the age of 2;0 never provides any 2sg forms. The main child, Marilena, showed high fluctuation of the use of 2sg, as she does not distinguish between 2sg and 3sg, especially only when the 2sg is formed with the addition of a final *-s* e.g. *graf-i-s* vs. *graf-i* whereas the other children of the study substitute the 3sg with the 2sg in all cases (Katis, 1984). Katis, (ibid.) explained the observed patterns phonemically; that is, while 3sg is phonemically unmarked, 1sg is formed with the use of what is referred as a stronger vowel, whereas 2sg is formed only with the addition of a final *-s* to the 3sg form. Katis (ibid.) then claimed that Marilena's failure to distinguish between 2sg and 3sg is due to her preference for open-syllable endings rather than a deficit in her grammatical knowledge (MG words are generally formed with an open-syllable ending). To further this claim, Katis, also discussed Marilena's

constant use of the alternative open-syllable ending for the 3pl form e.g. *dinun-e* instead of *dinun*.

Katis (ibid.) also observed that when person morphology is not missing or substituted, personal pronouns are provided or contextual information indicates which person was intended in the utterance.

Children in Katis' (ibid.) study did not show any problems with Number morphology; again, the information related to number is provided in the context.

Katis (ibid.) proposed that the categories of person and number are the first, of the verbal morphology, to be acquired with productive use already from the age of 2;6 even though with some errors in their production. Number and person errors are claimed to disappear only after the age of 2;9 (Katis, ibid.)

Katis (1984) proposed that the early acquisition of person and number categories, as opposed to tense, aspect, and mood, can be explained by the following two attributes of those categories:

- a) Person and number markers are perfectly regular, segmentable, word final markers and;
- b) They do not mark complex notions. The fact that person and number marker meanings are given through pronouns and noun inflections quite early seems to support the latter claim.

(Katis, 1984, p. 99)

6.5.2 Tsimpli

Tsimpli (1992/1996), as discussed in previous chapters, proposed the prefunctional stage, during which children's language lacks functional categories.

For the agreement category, Tsimpli (ibid.) claimed that at the prefunctional stage children use mostly the 3rd singular person for most of the verbs they produce. Katis' (1984) first developmental stages, discussed above, comply with Tsimpli's suggestion. Tsimpli (ibid.), within the prefunctional stage framework, proposed this sole use of 3sg to be instantiated because of unspecified agreement features in children's development, and therefore, lack of syntactic properties related to the agreement phrase. Tsimpli (ibid.) suggested that children use 3sg because of its unmarked nature, for instance 3sg is used in adult Greek, amongst other languages, in impersonal constructions where no thematic subject is required.

6.5.3 Christofidou and Stephany (2003)

Stephany (1985, cited in Christofidou and Stephany, 2003), within the pre- and proto-morphological framework, suggested that the lack of contrasting inflectional forms of the same verb in children's early speech defines the pre-inflectional stage. Stephany (ibid.) claimed that children may use inflected forms according to the adult model, however, the functional features of these forms are not formally distinguished. According to Stephany, the transition from the pre-inflectional to the inflectional stage is defined by the increase in grammaticisation, namely, children express different functions using different forms, and these forms can also express further functions, such as aspect, tense, person, and number, in a more differentiated way (Stephany, 1985, p. 225, cited in Christofidou and Stephany, 2003).

Stephany (ibid.) proposed that children's morphological development has two characteristics: it is piecemeal because acquisition is not achieved at once but on an item-by-item basis and systematic at the same time (systematicity is inherent in the grammatical categories the children choose from the input already in the very beginning

of development). Children initially use unmarked inflectional combinations and only later, at the proto-morphological stage, their development becomes systematic with the use of paradigmatic patterns.

Christofidou and Stephany (2003) reported on two databases, i.e. a longitudinal corpus of Christos from 1;7.11 to 2;1.27 and a selection from Mairi's speech obtained from the Stephany corpus (available in CHILDES, MacWhinney, 2000) from 1;9.17 to 1;9.29.

According to Christofidou and Stephany (*ibid.*), Mairi's speech can be already positioned in the initial phases of the inflectional stage. Christos at the ages of 1;8 and 1;9 uses two and three different verbs with clear inflectional marking, respectively and hence (based on the productivity criteria discussed later in this chapter), Christofidou and Stephany (2003) positioned his verbal and inflectional development at the pre-morphological stage (1;7 – 1;10). The onset of the proto-morphological stage for Christos' inflectional development, is placed at the age of 1;11 (Christofidou and Stephany, *ibid.*).

Christofidou and Stephany (2003) suggested that the rote learning forms used during the pre-morphological stage are the 'precursors' of grammatical rules and inflectional patterns that will emerge later. According to Christofidou and Stephany (2003, p. 117), the patterning of inflectional forms is the most important achievement in the development of Greek verbal inflection and shows that the children have entered the proto-morphological stage.

6.6 Analyses

6.6.1 Person morphology

The following section examines the accuracy of person morphology in Maria's and Eve's data. Two analyses were conducted: one to look at errors of form and a second one at errors of context. Subject-verb agreement accuracy was established by examining the person and number of the subject. This was defined from the context because in most utterances the subject is not overtly realised.

Tables 6.2 and 6.3 show the accuracy of singular and plural person morphology in Maria's speech. Correct use of person and number morphology is given with percentages and raw numbers (correct instances out of the total). Incorrect use contains only agreement errors. The MLU is given next to the biological age for ease of comparison.

Table 6.2: Maria's accuracy of person morphology of the verbal paradigm in the singular (errors of form)

| Maria Age | MLU | 1SG | | 2SG | | 3SG | |
|----------------------------|------------|------------|---------|------------|---------|------------|---------|
| | | | | | | | |
| 2;0 | 3.8 | 100% | n=12/12 | 100% | n=3/3 | 85.7% | n=18/21 |
| 2;2 | 3.4 | 100% | n=34/34 | 100% | n=27/27 | 92.3% | n=12/13 |
| 2;3 | 3.8 | 97.8% | n=45/46 | 100% | n=17/17 | 93.8% | n=61/65 |
| 2;5,4 | 4.2 | 100% | n=39/39 | 100% | n=19/19 | 98.5% | n=66/67 |
| 2;5,24 | 3.6 | 100% | n=33/33 | 100% | n=13/13 | 100% | n=66/66 |
| 2;7 | 3.7 | 100% | n=29/29 | 100% | n=10/10 | 97.7% | n=43/44 |
| 2;8,3 | 3.6 | 100% | n=29/29 | 100% | n=22/22 | 87.5% | n=14/16 |
| 2;8,27 | 4.0 | 100% | n=41/41 | 100% | n=9/9 | 100% | n=65/65 |

Table 6.2 shows that the use of singular morphology occurs from the onset of Maria's speech production and increases with time. At the age of 2;0 there are 12 correct instances of 1st singular, 3 correct instances of 2nd singular and 18 correct instances of 3rd singular in Maria's speech. The accuracy rate is also very high from the onset of the data collection, i.e. 100% of the 1st singular, 100% of the 2nd singular, and 85.7% of the 3rd singular suffixes have been used in the appropriate context. Most errors are observed with the 3rd singular person. The numbers and accuracy rate increase with age and with increase of MLU.

Table 6.3: Maria's accuracy of person morphology of the verbal paradigm in the plural
(errors of form)

| Maria Age | MLU | 1PL | | 2PL | | 3PL | |
|--------------|-----|------|---------|------|-------|------|---------|
| | | % | n | % | n | % | n |
| 2;0 | 3.8 | 100% | n=6/6 | 0% | n=0 | 100% | n=1/1 |
| 2;2 | 3.4 | 100% | n=7/7 | 0% | n=0 | 100% | n=4/4 |
| 2;3 | 3.8 | 100% | n=10/10 | 100% | n=1/1 | 100% | n=6/6 |
| 2;5,4 | 4.2 | 100% | n=15/15 | 100% | n=1/1 | 100% | n=16/16 |
| 2;5,24 | 3.6 | 100% | n=5/5 | 0% | n=0 | 100% | n=10/10 |
| 2;7 | 3.7 | 100% | n=3/3 | 100% | n=1/1 | 100% | n=8/8 |
| 2;8,3 | 3.6 | 50% | n=1/2 | 0% | n=0 | 100% | n=3/3 |
| 2;8,27 | 4.0 | 0% | n=0 | 100% | n=1/1 | 67% | n=2/3 |

Table 6.3 shows that plural marking can also be observed from the onset of the data collection. At the age of 2;0, there are 6 correct instances of 1st plural and 1 correct instance of 3rd plural. 2nd plural emerges at the age of 2;3. There is considerably less context for 2nd plural in the data. The accuracy rate is also very high from the beginning, i.e. 100% of 1st plural and 100% of 3rd plural at the age of 2;0. The use of plural morphology is limited compared to the singular one. There are substantially less contexts for plural. The number of use increases with age and although it appears to fluctuate, this is due to the length of each recording. Accuracy also remains very high throughout the data collection.

Examples of correct use of person and number morphology from Maria's data are outlined below:

98. Pano selis?

Pano **want/2sg**

Pano do you want?

(Maria 2;0)

99. pai alo sa pume, na pume alo

gone other will **say/1pl** toSUBJ **say/1pl** other

it's gone; we will say another one, let's say another one

(Maria 2;0)

100. afu thes na to valis ato gia na kimithune

since **want/2sg** toSUBJ it **put/2sg** this for toSUBJ **sleep/1pl**

Since you want to put this so we can sleep

(Maria 2;2)

101. xxx tora pezo xxx ato

xxx now **play/1sg** xxx this

I now play this

(Maria 2;2)

102. tha to valo na perpatisi

will it **put/1sg** toSUBJ **walk/3sg**

I will put/make it to walk

(Maria 2;3)

103. jaja, min kanis etsi, jaja
 granny, not **do/2sg** like-this granny
Granny, don't react this way, granny.
 (Maria 2;3)
104. xxx kanate, ntei, ntei
 xxx **do/2pl** speech particle
You did horsy-riding
 (Maria 2;5,4)
105. ta pedja vafune
 the children **paint/3pl**
The children are paining
 (Maria 2;5,4)
106. den pepatai ato
 not **walk/3sg** this
This doesn't walk
 (Maria 2;5,24)
107. xxx eki vjikane t'aterakja
 xxx there **came-out/3pl** the starsDIM
The stars came out (appeared) over there
 (Maria 2;5,24)

108. djavazo djafimisis
read/1sg advertisements
I read advertisements
(Maria 2;7)
109. ine to idjo onoma
is/3sg the same name
It's the same name
(Maria 2;7)
110. den kimomaste etsi
not **sleep/1pl** this-way
We don't sleep this way
(Maria 2;8,3)
111. ego tha sikotho apo ti thesi mu
I will **stand-up/1sg** from the seat my
I will stand up from my seat
(Maria 2;8,3)
112. xxx to pirane mech(r)i na ste(g)nosi ti petseta
xxx it **took/3pl** till toSUBJ **dry/3sg** the towel
They took it, till the towel was dried
(Maria 2;8,27)

113. den bori na to pjasi
not **can/3sg** toSUBJ it **catch/3sg**

He/she cannot catch it

(Maria 2;8,27)

Tables 6.4 and 6.5 present the accuracy rate of singular and plural person morphology in Eve's speech.

Table 6.4: Eve's accuracy of person morphology of the verbal paradigm in the singular
(errors of form)

| Eve Age | MLU | 1SG | | 2SG | | 3SG | |
|------------|-----|----------|-----------|----------|---------|----------|-----------|
| | | Accuracy | n | Accuracy | n | Accuracy | n |
| 1;7 | 2.0 | 100% | n=11/11 | 100% | n=26/26 | 100% | n=38/38 |
| 1;8 | 1.6 | 100% | n=25/25 | 100% | n=4/4 | 40% | n=2/5 |
| 1;9 | 2.0 | 98.2% | n=57/58 | 100% | n=39/39 | 98.4% | n=65/66 |
| 1;10 | 1.9 | 100% | n=93/93 | 100% | n=67/67 | 97.1% | n=104/107 |
| 1;11 | 2.3 | 100% | n=54/54 | 100% | n=34/34 | 100% | n=83/83 |
| 2;0 | 2.4 | 100% | n=66/66 | 98.6% | n=72/73 | 97.3% | n=185/190 |
| 2;2 | 2.5 | 100% | n=111/111 | 100% | n=55/55 | 100% | n=77/77 |
| 2;3 | 2.3 | 100% | n=51/51 | 100% | n=56/56 | 100% | n=159/159 |
| 2;5 | 2.8 | 100% | n=85/85 | 100% | n=73/73 | 100% | n=152/152 |
| 2;6 | 2.5 | 100% | n=83/83 | 100% | n=31/31 | 100% | n=80/80 |
| 2;9 | 2.8 | 100% | n=29/29 | 100% | n=18/18 | 100% | n=16/16 |
| 2;10 | 2.7 | 97.6% | n=41/42 | 100% | n=27/27 | 100% | n=37/37 |
| 2;11 | 4.4 | 100% | n=1/1 | 100% | n=9/9 | 100% | n=14/14 |

Table 6.4 shows that Eve has been using 1st, 2nd, and 3rd person singular suffixes from the onset of the data collection. At the age of 1;7 and with MLU of 2.0, there are 11 correct instances of 1st sg, 26 correct instances of 2nd sg, and 38 correct instances of 3rd sg suffixes. The accuracy rate is also very high from the onset of the data collection. 100% of the 1st, 2nd and 3rd singular suffixes have been used in the appropriate context.

Most errors involve the use of 3rd singular person. The use of singular morphology increases with time and accuracy rates remain very high throughout the data collection.

Table 6.5: Eve's accuracy of person morphology of the verbal paradigm in the plural (errors of form)

| Eve Age | MLU | 1PL | | 2PL | | 3PL | |
|------------|-----|----------|---------|----------|-------|----------|---------|
| | | Accuracy | n | Accuracy | n | Accuracy | n |
| 1;7 | 2.0 | 100% | n=4/4 | 0% | n=0 | 100% | n=1/1 |
| 1;8 | 1.6 | 0% | n=0 | 0% | n=0 | 0% | n=0 |
| 1;9 | 2.0 | 100% | n=18/18 | 0% | n=0 | 100% | n=3/3 |
| 1;10 | 1.9 | 100% | n=7/7 | 100% | n=2/2 | 100% | n=12/12 |
| 1;11 | 2.3 | 100% | n=6/6 | 100% | n=1/1 | 100% | n=20/20 |
| 2;0 | 2.4 | 100% | n=20/20 | 0% | n=0 | 100% | n=58/58 |
| 2;2 | 2.5 | 100% | n=25/25 | 0% | n=0 | 100% | n=5/5 |
| 2;3 | 2.3 | 100% | n=10/10 | 0% | n=0 | 100% | n=6/6 |
| 2;5 | 2.8 | 100% | n=13/13 | 100% | n=1/1 | 100% | n=8/8 |
| 2;6 | 2.5 | 100% | n=15/15 | 0% | n=0 | 100% | n=2/2 |
| 2;9 | 2.8 | 100% | n=4/4 | 100% | n=1/1 | 100% | n=3/3 |
| 2;10 | 2.7 | 100% | n=3/43 | 100% | n=3/3 | 100% | n=9/9 |
| 2;11 | 4.4 | 0% | n=0 | 0% | n=0 | 100% | n=3/3 |

Table 6.5 shows that plural morphology is also used from the beginning of the data collection. At the age of 1;7 there are 4 correct instances of 1st plural and 1 correct instance of 3rd plural in Eve's speech. The 2nd plural emerges later, at the age of 1;10.

There are substantially less contexts for plural in Eve's data. However, plural marking emerges from the first recording and the accuracy is very high.

Examples of correct use of person and number morphology from Eve's data are outlined below:

114. mole, tha paro go ti tada, ki tha figo
[discourse marker], will **take/1sg** I the handbag, and will **leave/1sg**
I will take the handbag and will leave
(Eve 1;7)

115. tha dosume sti mama
will **give/1pl** to-the mummy
We will give to mummy
(Eve 1;7)

116. kita to nini kimate
look the baby **sleep/3sg**
look the baby sleeps
(Eve 1;9)

117. xxx xxx i mamaka, o babakas de fovate
xxx xxx the mummyDIM, the daddyDIM not **afraid/3sg**
mummy, daddy is not afraid
(Eve 1;9)

118. mu klini to mataki
to me **close/3sg** the eyeDIM
he/she winks at me lit: he/she closes their little eye to me
(Eve 1;9)
119. ata kila petas sto tzaki
those woods **throw/2sg** into-the fireplace
you throw in the fireplace those woods
(Eve 1;9)
120. de to pilazume to tak [?] xxx
not it **touch/1pl** the tak xxx
we don't touch it, the tak
(Eve 1;10)
121. sfilizi o papus
whistle/3sg the granddad
the granddad whistles
(Eve 1;10)
122. pedja a pate xxx to xxx
children to **go/2pl** xxx the xxx
children you should go xxx the xxx
(Eve 1;11)
123. atos o athropos pu kathete xxx to tapezi

this the man which **sit/3sg** xxx the table

this man, the one sitting at the table

(Eve 1;11)

124. thelo na ftjatsis piti xxx

want/1sg toSUBJ **make/2sg** house xxx

I want you to make a house

(Eve 2;0)

125. thelune folitsa

want/3pl nestDIM

they want a little nest

(Eve 2;0)

126. theli na ksaplosi

want/3sg toSUBJ **lie-down/3sg**

he/she wants to have a lie down

(Eve 2;2)

127. pame na parume kati pu thelo

go/1pl toSUBJ **take/1pl** something that **want/1sg**

let's go to take something I want

(Eve 2;2)

128. na thimithume na parume eki kato

toSUBJ **remember/1pl** toSUBJ **take/1pl** there down
we need to remember to take down there

(Eve 2;3)

129. ke xxx pigane sto dasos
and xxx **go/3pl** in-the forest
and they went in the forest

(Eve 2;3)

130. apano sta dedra chothike
on-top in-the trees **hide/3sg**
he got hidden up there in the trees

(Eve 2;5)

131. na kanete tsulithra.
toSUBJ **do/2pl** slide
you should do the slide

(Eve 2;9)

The examples above show correct use of person and number marking in Maria's and Eve's language production. The examples contain a variety of verbal constructions with number and person marking highlighted.

To summarise, the rates of agreement error are very low in both Maria's and Eve's speech. 3rd singular person provides the context for most errors both in Maria's and Eve's speech. Error rates decrease with age. Both singular and plural morphology

are present and produced in Maria's and Eve's speech from the onset of the data collection. Plural however appears to have limited contexts in both datasets compared to singular.

6.6.2 Error analysis

The following section examines the distribution of agreement errors in Maria's and Eve's corpora.

Agreement errors consist of a substitution of one person and number verbal suffix for another.

Tables 6.6 and 6.7 present agreement errors in Maria's and Eve's data respectively. The Tables contain only⁴⁸ agreement errors.

Table 6.6: Maria's Agreement Errors

| Maria Age | MLU | Agreement Errors | Correct verbs |
|------------------|------------|-------------------------|----------------------|
| 2;0 | 3.8 | 7.9% (n=3) | 92.1% (n=35) |
| 2;2 | 3.4 | 1.3% (n=1) | 98.7% (n=80) |
| 2;3 | 3.8 | 3.6% (n=5) | 96.4% (n=136) |
| 2;5,4 | 4.2 | 0.7% (n=1) | 99.3% (n=147) |
| 2;5,24 | 3.6 | 0% (n=0) | 100% (n=122) |
| 2;7 | 3.7 | 1.1% (n=1) | 98.9% (n=92) |
| 2;8,3 | 3.6 | 4.5% (n=3) | 95.5% (n=64) |
| 2;8,27 | 4 | 0.9% (n=1) | 99.1% (n=118) |
| Total | | 1.9% (n=15) | 98.1% (n=794) |

⁴⁸ For a detailed analysis of all errors of both corpora see chapter 5.

In Table 6.6, the distribution of agreement errors is given for Maria’s entire data. Maria’s agreement error rate is at its highest at the beginning of the data collection at the age of 2;0, namely 7.9%. Between the ages of 2;2 and 2;7, agreement error rates fluctuate between 0.7% and 3.6%, therefore there is a decrease of errors with age. At 2;8,3 the error rate rises up to 4.5% and then during the last recording at the age of 2;8,27 falls to 0.9%. Maria’s mean agreement error rate is at 1.9%.

Table 6.7: Eve’s Agreement Errors

| Eve Age | MLU | Agreement Errors | Correct verbs |
|----------------|------------|-------------------------|-----------------------|
| 1;7 | 2 | 0% (n=0) | 100% (n=72) |
| 1;8 | 1.6 | 7.3% (n=2) | 93.7% (n=30) |
| 1;9 | 2 | 1.2% (n=2) | 98.8% (n=173) |
| 1;10 | 1.9 | 1.1% (n=3) | 98.9% (n=272) |
| 1;11 | 2.3 | 0% (n=0) | 100% (n=196) |
| 2;0 | 2.4 | 1.5% (n=6) | 98.5% (n=394) |
| 2;2 | 2.5 | 0% (n=0) | 100% (n=267) |
| 2;3 | 2.3 | 0% (n=0) | 100% (n=281) |
| 2;5 | 2.8 | 0% (n=0) | 100% (n=327) |
| 2;6 | 2.5 | 0% (n=0) | 100% (n=209) |
| 2;9 | 2.8 | 0% (n=0) | 100% (n=71) |
| 2;10 | 2.7 | 0.9% (n=1) | 99.1% (n=119) |
| 2;11 | 4.4 | 0% (n=0) | 100% (n=27) |
| Total | | 0.7% (n=14) | 99.3% (n=2438) |

Table 6.7 shows the agreement errors in Eve's data. The agreement error rate is at its highest at the age of 1;8, i.e. 7.3% and thereafter the rate decreases with age, namely between 1.1% and 1.5% at the ages of 1;9 to 2;0. At 2;10 the agreement error rate is at its lowest at 0.9%. Eve's average agreement error rate is 0.7%.

6.6.3 Error analysis by context

This part presents the second analysis that looks at the breakdown of errors from a context point of view. Tables 6.8 and 6.9 below show agreement errors in Maria's and Eve's data presented as of their context, i.e. what form was used and what was the required context.

Table 6.8: Maria's distribution of Agreement errors (context errors)

| Form used | Form required by context | | | | | |
|-----------|--------------------------|-----|-----|-----|-----|-----|
| | 1sg | 2sg | 3sg | 1pl | 2pl | 3pl |
| 1sg | | | 1 | | | |
| 2sg | | | | | | |
| 3sg | 7 | 5 | | | | |
| 1pl | | | | | | 1 |
| 2pl | | | | | | |
| 3pl | | | 1 | | | |

Table 6.8 shows that 14 out of 15 of Maria's errors (93%) occur in singular contexts, and from those, most involve 3rd singular person. The majority of 3sg errors

involve a 3sg used instead of 1sg (n=7/15). This phenomenon is discussed in the following section. The second most frequent error (n=5/15) involves the use of 3sg instead of 2sg. There is only one case of number error i.e. a 3pl used instead of a 3sg.

Table 6.9: Eve's distribution of Agreement errors (context errors)

| | Form required by context | | | | | |
|-----------|--------------------------|-----|-----|-----|-----|-----|
| Form used | 1sg | 2sg | 3sg | 1pl | 2pl | 3pl |
| 1sg | | | 2 | | | |
| 2sg | | | | | | |
| 3sg | 8 | 1 | | 1 | | 2 |
| 1pl | | | | | | |
| 2pl | | | | | | |
| 3pl | | | | | | |

Table 6.9 shows that Eve's agreement errors are similar to Maria's, namely 11 out of 14 (78%) occur in singular contexts. A similar pattern is shown for the 3rd singular person involving the majority of errors occurred, where a 3sg is used in 1sg and 2sg contexts. There are three number errors where singular was used instead of plural all involving 3sg person.

6.6.3.1 *Maria's errors*

As discussed previously in the error analysis by context, Maria's agreement errors can be categorised in two types: a) use of a 1st person singular in a 2nd and 3rd person

singular context as shown in (132) – (134) below and b) an agreement mismatch between the main and the embedded clause, as given in (135) below:

132. atos echo (target: aftos echi)
he **have/1sg**
he, I have (instead of *I have*)
(Maria 2;2)
133. na zosi ato (target: na dosis afto)
toSUBJ **give/3sg** this
he/she should give this (instead of *you should give this*)
(Maria 2;0)
134. n' afisi kato (target: n' afisis kato)
toSUBJ **leave-let/3sg** down
he/she should leave/let down (instead of *you should leave/let this*)
(Maria 2;0)
135. de boruse na to kano ato to baloni (target: de boruse na to kani)
not **can/3sg** to it **do/1sg** this the balloon
He was not able I to do this balloon (instead of *He was not able to do this ballon*)
(Maria 2;7)

In (132) Maria uses 1st person singular in a 3rd singular context. Examples (133) and (134) are substitution errors involving the use of 3sg instead of 2sg person. In

Maria's data this is the second most frequent error (n=5/15). In (135) the verb of the main clause in 3rd singular is followed by a 1sg subjunctive/dependent verb which does not agree with the main verb. This type of error might also indicate a gap in the child's knowledge of the verbs' subcategorisation frame i.e. the type and arguments of the verb. For instance if the verb *thelo* (= I want) was used instead as the main verb in the same sentence as in (136) below, then the sentence is not ungrammatical:

136. den ithele na to kano ...

Not **want/3sg** to it **do/1sg**...

He didn't want me to do this...

6.6.3.2 Eve's errors

Based on the analysis of errors by context, there are three main error types identified in Eve's data: a) use of a singular 3rd person morpheme in a plural context, as shown in examples (137) and (138) below, b) use of a 3rd person morpheme in 1st person context, as shown in the examples (139) below, and c) substitution errors involving the use of 3sg instead of 2sg person as shown in example (140).

137. theli tsulit(h)ra ta pedja (target: thelun tsulithra ta pedja)

want/3sg slide the kidsPL

the kids wants to go to the slide (instead of the kids want to go to the slide)

(Eve 2;0)

138. theli na kathisi ta pedja (target: thelun na kathisun ta pedja)

want/3sg to sit the kidsPL

the kids wants to sit down (instead of the kids want to sti down)

(Eve 2;0)

139. ego [*] ftjatsi [?] xxx to xxx (target: ego na/tha ftjakso)

I1sg 0prt **make/3sg** xxx the xxx

I will makes the xxx (instead of I will make the xxx)

(Eve 2;0)

140. n' anitsi (target: n'aniksis)

toSUBJ **open/3sg**

you he-should open (it) (instead of you should open it)

(Eve 1;10)

In (137) and (138) the 3sg verb is used with a subject in the plural; these are both number (and person) errors. In example (139) Eve uses the personal pronoun *I* with a 3rd singular verb; this is a person (not number) error. Example (140) is the only instance in Eve's data where a 3sg person is used in a 2sg context.

To summarise, both Maria's and Eve's agreement error rates are low and accuracy of agreement morphology increases with age. The examples given above, show correct and incorrect use of agreement morphology. Children at an early age of their language development appear to commit a frequent substitution error when they use 3rd singular verbs instead of 1st and 2nd singular.

Tsimpli (1992), as discussed at the beginning of this chapter, suggested that children within the prefunctional stage use, predominantly, verbs in 3rd singular person due to unspecified agreement features in children's grammar. A similar pattern is

observed in Katis' study (1984) where children do not distinguish between the 2sg and 3sg. Katis (ibid.) claims that this is due to the fact the 2sg is phonemically more marked than 3sg because of the addition of the final *-s*, therefore children prefer an open syllable ending rather than the ones with an additional ending. Maria's and Eve's agreement errors are far too limited to be considered to be within the prefunctional stage, however, such limited presence of these types of errors, may provide evidence for a prefunctional stage.

6.6.3.3 *Inappropriate use of 3rd singular person – self reference*

Guasti (1994) maintained that children often refer to themselves with their own name in 3rd singular, an observable fact in their parents' speech too (Child Oriented Speech). This inappropriate use of agreement involves the use of a 3rd person singular in a 1st person singular context, as shown in the dialogues given below between children and adults:

141. MAR: echis ki esi mustakja.

You have moustaches too

ADU: esi echis?

Do you have any?

MAR: zen echo.

I don't (have)

MAR: na va(l)is ke ati mutakja.

You should put to her moustaches too

MAR: echi k'i Maria.

Maria has too

ADU: k'i Maria echi mustakja.

Maria has moustaches too

(Maria, 2;2)

142. ADU: a re pjos kimate mesa re?

who sleeps inside there?

EVE: i Eva.

Eva

ADU: pja Eva?

Eva who?

(Eve 2;5)

The dialogues in (141) and (142) above between Maria and Eve respectively and adults show how Maria and Eve respond to the adults' use of 3rd singular when referring to them; they both use 3rd singular too, referring to themselves.

Also, the examples below are utterances produced by Eve's mother and granmother during the tape recordings where the adults refer to Eve using the 3rd singular person:

143. tha feri sokolata tis Evis?

will **bring/3sg** chocolate the Eva?

will he/she bring a chocolate bar for Eva?

(Eve 1;10)

144. ti kani i Eva agapi mu?

What **do/3sg** the Eva darling my?

What does Eva do, my darling?

(Eve 1;7)

145. tha ta riksi kato i Eva

will them **throw/3sg** down the Eva

Eva will throw them down

(Eve 1;9)

Examples of utterances of Maria and Eve referring to themselves in 3rd singular rather than 1st singular are given below:

146. ze seli

not **want/3sg**

(She) doesn't want

(Maria, 2;0)

147. de lei

not **tell/3sg**

(She) doesn't tell

(Maria, 2;2)

148. ine zio minon

is/3sg two months

(She/he) is two months (old)

(Maria, 2;2)

149. o athropos kani kako (s)ti Maria.
the man **do-hurt/3sg** the Maria
The man hurts Maria

(Maria, 2;3)

150. i Eva kunja bela
the Eva **swing/3sg**
Eva will do the swing

(Eve 1;7)

151. i Eva pi je xxx ego
the Eva **went/3sg** xxx I
Eva went xxx I

(Eve 2;0)

The examples above (146) – (151) show Maria and Eve using the 3rd singular person rather than the 1sg, when referring to themselves, following the example of the adult speech.

As shown in the examples above (141) – (151) Maria and Eve use the 3rd singular person rather than the 1st singular, when referring to themselves. Such inappropriate use is often encountered in children's early speech and it may be considered to be the result of children and often their parents referring to them with their own name (Guasti, 1994).

6.7 Productivity

6.7.1 *Productivity of the use of person and number marking*

The following part provides a qualitative analysis of Maria's and Eve's use of person and number morphology. The qualitative analysis was conducted to address the issue of productivity of the use of person and number suffixes. This analysis will investigate whether each suffix is being used only with one verb in each recording, or whether it is being used with a variety of different verbs. In addition, a second analysis was conducted to examine whether each verb is being used with only one suffix or with a variety of different suffixes from the inflectional paradigm.

This methodology was based on Pizzuto and Caselli's (1992) proposed criteria of morphological productivity, according to which a given inflectional form is considered to be productive when either or both of the following conditions hold:

- a. The same verb root appears in at least two distinct inflected forms and
- b. The same inflection is used with at least two different verbs

(Pizzuto and Caselli, 1992)

Guasti's (1994), productivity criteria are in agreement with Pizzuto and Caselli's. Guasti, (ibid.) suggested that children from a very early age use a variety of agreement morphemes. Guasti (ibid.), in her analysis of data from Martina and other Italian-speaking children, claimed that Italian speaking children use different verbs with the same agreement morphemes, for example, they use the third person suffix *-e* with different roots:

cad-*e* fall/3sg
piang-*e* cry/3sg

(Martina 1;8 in Guasti, 1994)

Guasti (ibid.) also suggested that early Italian-speaking children use the three singular agreement markers with the same verbal root:

mett-*o* put/1sg
mett-*i* put/2sg
mett-*e* put/3sg

(Diana 1;11 in Guasti, 1994)

The examples presented above from Guasti, cover Pizzuto and Caselli's minimal requirements of inflectional development.

Not far from what is proposed so far and within the pre- and proto-morphological framework, Bittner, Dressler and Kilani-Shock (2000) proposed the mini-paradigm criterion of inflectional productivity defined as:

... a non isolated set of minimally three accurate and distinct inflectional forms of the same verbal lexeme produced spontaneously in contrasting contexts.

(Bittner, Dressler and Kilani-Shock, 2000, p.5)

Compatible with the pre/proto-morphological approach, the Percentage of Base Form (PBF), proposed by Voeikova and Gagarina (2002) is a measure for determining

early inflectional development even before the onset of mini-paradigms. PBF measures the base forms occurring in children’s utterances and determines the percentage of tokens of unmarked inflectional forms (Stephany and Voeikova, 2009). The base form is the default (unmarked) form used by the children that usually corresponds to the citation form (bare stem) in the adult language (Voeikova and Gagarina, *ibid.*). This needs to be examined on a language by language basis to capture language specific differences. Voeikova and Gagarina (*ibid.*) propose that the PBF gradually decreases as the child’s development progresses.

The analyses for productivity discussed in the following section, are based solely on Pizzuto and Caselli’s and Guasti’s productivity criteria.

6.7.2 *Productivity in Maria’s and Eve’s speech*

Starting with Pizzuto and Caselli’s (*ibid.*) and Guasti’s (*ibid.*) criteria for productivity this section will discuss the productivity of Maria’s and Eve’s data. Similar examples given in Guasti (1994) are observed in Maria’s and Eve’s corpus. The examples below show that the 1st singular person suffix *-o* is used with a variety of different verbs from the onset of the data collection, i.e. 2;0 for Maria and 1;7 for Eve:

152. oochi de sel-*o* ato
 no not **want/1sg** this
 I don’t want this

(Maria, 2;0)

153. t’agap-*o* agapi mu
 it **love/1sg** darling my

I love it, my darling

(Maria, 2;0)

154. edo, se k(r)ata-*o* ego
here, youACC **hold/1sg** I
here, I am holding you

(Maria, 2;0)

155. de bor-*o*
not **can/1sg**
I cannot

(Eve 1;7)

156. to baz-*o* bano
it **put/1sg** up-there
I put it up there

(Eve 1;7)

157. t(r)o-*o* moni mu
eat/1sg alone mine
I eat all by myself

(Eve 1;7)

The three singular agreement markers are used by Maria and Eve with the same verbal root; in the examples given below, the verb *vazo/valo* (= put) of the 1st conjugation is used in the same recording, with all three singular agreement markers

with the perfective stem. These examples are not from the onset of the data collection, however, from very early age, i.e. 2;2 for Maria and 2;0 for Eve:

val-*o* put/1sgPERF
 val-*is* put/2sgPERF
 val-*i* puts/3sgPERF

158. na to valo eso ato
 toSUBJ it **put/1sg** out this
I am going to put this out

(Maria 2;2)

159. afu thes na to valis ato gia na kimithune
 since want2sg toSUBJ it **put/2sg** this for to sleep
Since you want to put this (to bed) so they can all sleep

(Maria 2;2)

160. tora (th)a vali ti Maria
 now will **put/3sg** the Maria
He/she will put Maria, now

(Maria 2;2)

161. the(1)o na valo ekino
 want1sg toSUBJ **put/1sg** that
I want to put that one

(Eve 2;0)

162. ato pu (tha) to valis?
This where it **put/2sg**
Where are you going to put this?

(Eve 2;0)

163. na va(1)i (s)piti
toSUBJ **put/3sg** house
he/she wants to put a house

(Eve 2;0)

At later recordings, i.e. Maria at 2;5,4 and Eve at 2;2, the verb *pao*⁴⁹ (= go) is used in all singular contexts.

164. pa-o go/1sg
pa-s go/2sg
pa-i goes/3sg

(Maria 2;5,4) (Eve 2;2)

According to Guasti (ibid.), productive use of agreement morphology is also observed with the copula *be* which presents irregular conjugation. In the examples given below the copula *be* is used in singular contexts with the suffixes of singular

⁴⁹ *Pao* is a single-stem verb i.e. does not present a morphological stem distinction between perfective and imperfective.

morphology. The examples in (165) – (170) are from more advanced recordings from the data collection, i.e. Maria at 2;5,24 and Eve 2;6:

165. oti ime kalo pedi, lei
that **am/1sg** good child, says3sg
He/she says, that I am a good child

(Maria 2;5,24)

166. ise kala Maria?
are/2sg alright Maria?
Are you alright Maria?

(Maria 2;5,24)

167. eki ine i komot(r)ia
there **is/3sg** the hairdresser
The hairdresser is there

(Maria 2;5,24)

168. ime apano alogo
am/1sg on horse
I am on the horse

(Eve 2;6)

169. ise skadalakos
are/2sg trouble-maker

You are a trouble-maker

(Eve 2;6)

170. ochi den ine o prigipas aftos
no not **is/3sg** the prince he
No, he is not the prince

(Eve 2;6)

The following examples (171) – (176) show the use of plural verbal morphology in Maria's and Eve's data. The earliest ages of plural morphology in the examples below are 2;3 for Maria and 1;10 for Eve:

171. xxx na pezume me ato
xxx toSUBJ **play/1pl** with this
(Let's/we will) play with this

(Maria 2;3)

172. ti sa selate?
what will **like/2pl**?
What would you like?

(Maria 2;3)

173. ta pedja vafune
the children **paint/3pl**
The children paint

(Maria, 2;5,4)

174. ela na to dume
come toSUBJ it **see/1pl**
Let's see it

(Eve 2;0)

175. ti me se(r)nete xxx?
what me **drag/2pl** xxx?
What do you drag me for?

(Eve, 1;10)

176. tha kolibis(une)
will **swim/3pl**
They will swim

(Eve 2;0)

The data presented above, suggest that Maria and Eve have started to build a verbal paradigm analysing the verb into its root and its inflectional affix often from the onset of the data collection if not from a very early age; in other words, children do not learn verbal forms as frozen expressions but apply a productive rule in using and producing them. Limited use of some components of the verbal paradigm, such as plural morphology, does not imply lack of knowledge or evidence that children do not know agreement.

6.7.3 *Frequency analysis*

The following analysis is set to examine the frequency of use of singular morphology in adults' speech and thereafter compare this to Maria's and Eve's data. This is to address question 8, repeated below for convenience:

8. Are more frequently used verb forms of person/number in adults more frequent in children too?

From a usage-based point of view the prediction for question 8 is that children use more frequently verb forms used frequently in the adult language they hear.

6.7.3.1 *Method*

The use of CLAN frequency programmes was employed for the analyses of the adults' speech in Maria's and Eve's recordings. The frequency analysis considered 1st, 2nd and 3rd singular endings of the first and second conjugation as follows:

1st and 2nd conjugation

| Person | Singular endings |
|-----------------------|-------------------------|
| 1st | <i>-o</i> |
| 2nd | <i>-is</i> |
| 3rd | <i>-i</i> |

All verbs with the above endings were considered regardless their aspect distinction, i.e. perfective or imperfective. The aim of this analysis was to compare the use of singular inflectional endings in adults' language to Maria's and Eve's speech

production. I have selected only singular morphology since the use of plural morphology in Maria's and Eve's speech is limited compared to the singular one, although productively used.

6.7.3.2 Results

The results of the frequency analysis are presented in the following tables:

Table 6.10: Maria's and adult use of singular inflectional endings

| | 1st singular | 2nd singular | 3rd singular | Total |
|--------|--------------------------------|--------------------------------|--------------------------------|--------------|
| Maria | 35.9% (n=252) | 16.2% (n=114) | 47.9% (n=334) | (n=700) |
| Adults | 19.9% (n=21) | 47.6% (n=50) | 32.5% (n=34) | (n=105) |

Table 6.11: Eve's and adult use of singular inflectional endings

| | 1st singular | 2nd singular | 3rd singular | Total |
|--------|--------------------------------|--------------------------------|--------------------------------|--------------|
| Eve | 31.3% (n=682) | 23.2% (n=506) | 45.5% (n=990) | (n=2178) |
| Adults | 19.2% (n=651) | 34.2% (n=1155) | 46.6% (n=1571) | (n=3377) |

Table 6.10 and 6.11 show the percentages and numbers of use of singular inflectional morphemes in Maria's, Eve's and the adult language from their respective datasets. Maria's and Eve's use of singular morphology is similar for each person i.e. the 3sg person has the highest rate of use (47.9% and 45.5% respectively), followed by the 1sg (35.9% and 31.1% respectively) and then by the 2sg (16.2% and 23.2%

respectively). The difference between Maria's and Eve's use of singular morphology is greater in the 2sg.

The adult use of singular morphology is different to Maria's data from Eve's data. In Maria's data the adult use of 1st singular is much lower than Maria's 19.9% vs. 35.9%, while the adult use of 2nd singular is much greater than Maria's, i.e. 47.6% vs. 16.2%. Finally, the adult use of 3rd singular is again lower than Maria's, namely, 32.5% vs. 47.9%. In Eve's data, similar to Maria's, the adult use of 1st singular is lower than Eve's, i.e. 19.2% vs. 31.3%, whereas the adult use of 2nd singular is again higher, namely, 34.2% vs. 23.2%. Finally, the adult use of 3rd singular is very similar to Eve's use, i.e. 45.5% vs. 46.6%. The adult speech production in Maria's dataset is very limited compared to the adult speech production in Eve's data, and therefore this may have an effect on the rate of use of singular morphology. These results will be discussed in details in the discussion section, at the end of this chapter.

6.8 Summary and discussion

To summarise so far, from the analyses presented in this chapter, the results suggest the Maria's and Eve's use of person and number morphology is most of the times used correctly with a very small rate of error. The accuracy of person and number marking morphological use was also checked by context and the results suggest high accuracy of the inflectional endings for person and number. Maria and Eve use number and person morphology productively, based on productivity criteria discussed above, repeated below for convenience:

- a. The same verb root appears in at least two distinct inflected forms and

- b. The same inflection is used with at least two different verbs

(Pizzuto and Caselli, 1992)

A frequency analysis, within the usage-based framework, suggests that Maria's and Eve's use of singular morphemes is very similar distributed amongst them. Maria's and Eve's use of singular morphology, however, is not similar to the adult use of singular.

The results of the analyses are compatible with the hypothesis that correct agreement features on verbal inflectional morphology are acquired very early in children's language development (cf. VEPS, VEKI and EMC).

The average rates of person and number errors in Maria's and Eve's data is 1.9% and 0.7% respectively. These results are also compatible with results on agreement error presented from other early languages (see Table 1). The average rate of agreement error in a variety of child languages is 4% and thus Maria's and Eve's rates are within this range. The results provide further support to the claim that correct agreement features of the verbal inflectional domain are available in children's grammar from early stages in the course of acquisition.

The patterns of emergence of person and number categories observed by Katis (1984) are different from the ones observed in Maria's and Eve's data. Katis (*ibid.*), as discussed at the beginning of the chapter, suggested that the order of emergence of inflectional markers is singular before plural and 3rd person before 1st and 2nd. Maria and Eve use all person and number markers from the onset of the data collection, i.e. 2;0 and 1;7 years of age respectively. Katis' (*ibid.*) findings showed a late emergence (after 2;0) of all persons except 3sg which is considered to be the unmarked one in the verbal paradigm. In Maria's and Eve's data all persons are used from an early age, i.e. the onset of the corpus collection (2;0 and 1;7 respectively); there is however a strong

preference for the 3rd person singular. The children in Katis' (ibid.) study used and were able to distinguish 3sg and 1sg and occasionally produced 2sg. The same patterns are observed in Maria's and Eve's speech. In terms of production and patterns observed, Maria's and Eve's speech production appears to be at a more advanced stage of development compared to the children analysed in Katis' study. In fact, Katis (ibid.) claims that Marilena, in the longitudinal data, although used from very early (age 2;6), number and person errors do not cease until the age of 2;9. This is much later compared to Maria's and Eve's production, as discussed in section 6.6.1 and section 6.6.2.

From a constructivist point of view (three stages in children's acquisition of inflectional morphology; the pre-morphological, the proto-morphological and the morphology proper stage (Dressler, 2003; Bittner et al., 2003), Maria's and Eve's speech appears to be at the third stage of development. Using productivity criteria (Guasti, 1994; Pizzuto and Caselli, 1992; Bittner, Dressler and Kilani-Shock, 2000) the findings of the present study suggest that Maria's and Eve's speech can be already positioned in the morphology proper stage, i.e. their morphological development and use and knowledge of inflectional markers resemble progress towards the adult model. These findings are different from Christofidou and Stephany's (2003) findings, according to which, Mairi's speech (age 1;9) is already at the inflectional stage whereas Christos' inflectional development is positioned in the pre-morphological stage (1;7 – 1;10).

The frequency analysis reveals that Maria and Eve use the singular morphemes in a parallel fashion, i.e. 3sg at a higher rate, then 1sg and then 2sg. These results are not similar to the adult use, as discussed previously, i.e. 1sg child>adult, 2sg child<adult, 3sg mixed: child≥adult. The nature of the data and the conversational context may denote that adults use the 2nd singular person more often than in other circumstances, i.e.

use of 2sg when addressing the children. Similarly, the higher use of 1sg in children's speech may, again, be the result of the discourse context in which children's responses are replies related to certain structures they hear. To conclude, the frequency analysis, suggests that Maria and Eve use of singular morphology is very similar, however, the adult use is not.

From a morphological point of view Xanthos et al. (2009) suggested that children exposed to morphologically rich languages acquire morphology quicker than children exposed to morphologically poor input and this happens via the input from the adults (child directed speech). This view is also reflected in Hyams' (2008) Stem Parameter Hypothesis, according to which if the inflectional component of a language is part of the core grammar of it, then acquisition of this component is easier, inflection being closely determined by UG parameters. Bearing in mind that Greek is a richly inflectional language, the findings of the analyses suggesting that agreement features are acquired early in Maria's and Eve's productions may reflect Xanthos' et al., (ibid.) and Hyams (ibid.) proposal of speedy and early acquisition of those features.

7 GENERAL DISCUSSION AND CONCLUSIONS

The aim of this chapter is to bring together the results of the analyses carried out in chapters 4, 5, and 6, provide an overview of the main findings from each chapter, and discuss them in relation to previous literature and language acquisition theories. The analyses were based on generative and usage-based approaches to language acquisition. Premises from both theories, related to the questions of this study, were addressed in the analyses.

The current thesis was set out to examine the development of some of the elements of the verbal domain in the spontaneous speech production of two Greek speaking children. These elements of the verbal domain are summarised in terms of their analyses related to the features in Table 2.1, repeated here as 7.1 for convenience:

Table 7.1: Analyses related to the features of the verbal domain in MG examined in this thesis

| Phenomenon | Category | Analyses | Chapter |
|-------------------|--------------------------|---|----------------|
| Tense | Past | regular sigmatic augment salience | §4 - (§5) |
| Inflection | Finiteness Inflection | modal particles aspect telicity | §5 |
| Agreement | Number Person | singular plural 1 st , 2 nd , 3 rd 3-suffix | §5 - §6 |

The research questions, results, and findings for each of the above verb analyses will be discussed in the following sections.

7.1 Past Tense

For the acquisition of past tense, the analyses were set out to examine two questions:

- a. From a generative point of view, this study investigated the difference between the sigmatic and non-sigmatic past tense in the children's production over time.
- b. From a usage-based point of view, this study examined if children's past tense forms can be traced back to the child's earlier production, to the adult's past tense forms, and which of the past verb forms are novel?

The first question examined the use of morphological features for the formation of sigmatic (ruled-based, regular) vs. non-sigmatic (allomorph, irregular) past tense and the mechanism of acquisition of these forms (single vs. dual mechanism). The predictions for this analysis are that sigmatic formation is acquired earlier than non-sigmatic one and used more productively being the regular formation for past tense. Within the same framework, a further analysis was performed to investigate the effect of salience in language acquisition, i.e. augmentation and regularity, as proposed by Mastropavlou (2006). The prediction of this analysis is that children will acquire the more salient past tense forms earlier than the less salient ones.

The analyses for the second question employed a trace-back procedure to test firstly if children use more often the past tense constructions that they used in their earlier production, and secondly, if these frequently used past tense constructions are frequently used in the adult input. The predictions for this question are that the frequent past tense forms can be traced back and they are used more often in the children's previous speech production and are frequent in the adult's language, suggesting that

frequent forms are acquired earlier than non-frequent ones. The analysis of frequency of past tense in the adult speech was based on the Hellenic National Corpus available from the Institute for Language and Speech Processing.

Previous research on the acquisition of sigmatic and non-sigmatic past tense formation in MG suggested that there is an asymmetry in the production and comprehension of sigmatic and non-sigmatic in children's early speech.

In Stathopoulou (2009), the results of a cross-sectional study of past tense showed that typically developing and Down's Syndrome children preferred past tense forms with sigmatic morphology to those with non-sigmatic one. Higher scores of both groups in choosing the sigmatic past in novel verbs supports the claim that the sigmatic past tense is rule-based, and thus, the most commonly used pattern in MG.

Stavrakaki and Clahsen (2009) also showed that there is an asymmetry between sigmatic and non-sigmatic morphology in typically developing children. The findings of Stavrakaki and Clahsen support the claim that the sigmatic past tense is the default formation in MG and are consistent with the dual-mechanism account of acquisition. From a developmental point of view, Stavrakaki and Clahsen (ibid.) interpret these results as suggesting that the acquisition of the non-sigmatic past tense takes longer than the sigmatic past tense.

The first analysis examined the sigmatic (regular) and non-sigmatic (irregular) formation of past tense in Maria's and Eve's data. The use of sigmatic vs. non sigmatic morphology varies over time: early on, children use more sigmatic than non-sigmatic forms. The use of non-sigmatic forms increases with time and in the final recordings the use of sigmatic past tense forms increases again. The results suggest that non-sigmatic forms (irregular formation) are used less often than sigmatic ones (regular formation), especially at the earliest stage of development.

These findings show an early asymmetry between sigmatic and non-sigmatic past tense. This can be considered to provide evidence and support for the dual mechanism of acquisition of past tense, since in Maria's and Eve's production the use of regular and irregular past tense morphology does not appear to be acquired in the same way from the onset of the data collection. These findings are in agreement with the results from previous studies, as discussed above (cf. Stathopoulou, 2009; Stavrakaki and Clahsen, 2009).

In relation to the single and dual mechanism hypothesis of acquisition, MG also displays regular and irregular morphology in its past tense system, i.e. the sigmatic and non-sigmatic past tense. The sigmatic form is considered to be the regular, rule-based form of the past, whereas, the non-sigmatic past tense forms are irregular and non rule-based. The sigmatic form is predictable and morphologically transparent while the non-sigmatic one is morphologically less transparent (Stavrakaki and Clahsen, 2009; Stathopoulou, 2009; Stathopoulou and Clahsen, 2010).

The low use of non-sigmatic past tense at the first recordings provides support for a dual mechanism of acquisition of past tense features. Non-sigmatic (irregular) past tense may be less frequently used, especially at early stages, because children learn and store these forms on a one-by-one basis as independent lexical entries, as opposed to regular past tense forms which are rule-based. The sigmatic past tense may be easier to acquire than the non-sigmatic one because it is rule-based and its formation is regular. The use of sigmatic and non-sigmatic past tense, however, varies over time and in later recordings the non-sigmatic one is more prominent. This finding may suggest that children from an early age also master the irregular formation of past tense.

In relation to past tense formation a further analysis considered a variety of features, i.e. augment, stress shift, allomorphy etc. These analyses were aimed at examining how salience affects language acquisition based on Mastropavlou's study.

Mastropavlou (2006) introduced the notion of *salience*. Past tense formation in MG is expressed through morphological and phonological features. These features are interpretable at the PF level and such interpretability relates to *salience* (Mastropavlou, *ibid.*). The use of these features defines the degree of salience in past formation. Mastropavlou (*ibid.*) proposes that [+augment] is more salient, whereas [-augment] is less salient and irregular past tense is more salient than regular past tense.

Mastropavlou (2006) studied the effect of salience in two groups of participants, i.e. children with SLI and typically developing children as the control group using examples of verbs fulfilling three salience conditions (discussed in section 4.5.4 but repeated here for convenience):

- | | |
|---|--|
| 1 st condition: Regulars +S: 18 items | e.g. ζωγραφίζω–ζωγράφισα (= <i>draw/drew</i>) |
| 2 nd condition: Regulars ++S: 12 items | e.g. γράφω – έγραψα (= <i>write/wrote</i>) |
| 3 rd condition: Irregulars: 10 items | e.g. πίνω – ήπινα (= <i>drink/drank</i>) |

(Mastropavlou, 2006: 74)

Mastropavlou's results suggested that there were differences among the three tested conditions, i.e. the 1st condition achieved the lowest scores for all groups while the 3rd condition achieved the highest scores. Therefore, both groups performed better with verbs of higher salience i.e. ++S, +Irr.

Mastropavlou (ibid.) proposed that before the age of four, salience has an effect in children's development (typically developing children); after the age of four, children grow out of this stage. Mastropavlou (ibid.) interprets these results as providing evidence for the dual mechanism, in the sense that there are two different routes for the acquisition of regular and irregular morphology.

The analysis of salience in this thesis considered the use of *augment* and *regularity* in sigmatic and non-sigmatic past tense in Maria's and Eve's data. The analysis revealed that the sigmatic past tense forms are mostly augmented and regular whereas the non-sigmatic past tense forms are non-augmented and irregular. Considering the results of the use of sigmatic and non-sigmatic past tense, discussed above, sigmatic past tense is used more often at the first recordings in Maria's and Eve's data, therefore, the earliest used past tense forms are the less salient ones, namely, 1st and 2nd conditions according to Mastropavlou's classification.

In relation to Mastropavlou's classification of salience, the findings of this analysis provide no support of the claim that [+salient] items are easier to acquire than [-salient].

In terms of development, the findings do not provide evidence of an early stage when children use only salient past tense and thereafter grow out of this stage. Maria and Eve use past tense regular and irregular morphology from the onset of the data collection, regardless of salience, although with some variation of use of the one over the other, overtime. Such variation of use may provide evidence for the dual mechanism, as discussed previously.

Based on usage-based studies, a trace-back procedure and a frequency analysis were performed in order to consider the effect of adult input frequency in Maria's and Eve's production of sigmatic and non-sigmatic past tense.

Usage-based studies claimed that frequency of the adult input has an effect on children's speech (Lieven, et al., 2003; Lieven et al., 2009; Dabrowska and Lieven, 2005; Lieven, 2006; Tomasello, 1992, 2000, 2003, 2006).

For MG, Stathopoulou, (2009) suggested that the results of her study (both DS and typically developing children preferred past forms with sigmatic morphology to those with non-sigmatic past constructions) could be due to frequency effects. Stavrakaki and Clahsen (2009) also observed that the sigmatic and non-sigmatic perfective past tenses have a different frequency distribution. Stavrakaki and Clahsen (ibid.) performed a frequency analysis in a large corpus of Greek words (Neurosoft Language Tools) and found that the vast majority of perfective past tense forms are sigmatic rather than non-sigmatic, i.e. 2119 sigmatic vs. 147 non-sigmatic past tense forms. The authors claim, that from a connectionist point of view, frequency may explain the prominence of the sigmatic past tense over the non-sigmatic.

The findings of the frequency analysis in Maria's and Eve's data revealed that children use frequently those verbs that are frequently used in the adult language (from the Hellenic National Corpus). Children also produce the verbs they know and used before by expanding on these constructions, as suggested by the trace-back procedure.

The findings of this analysis suggest that the adult input plays a major role in the speech production of children exposed to a given language and their grammatical development.

However, the trace-back procedure and frequency analysis also revealed that Maria and Eve use many novel verbs in their speech that cannot be traced-back in their previous production or in the adult input from the recordings. This provides counter-evidence to the usage-based claim that children acquire verbs on a verb-by-verb basis and produce new constructions building on their previous knowledge of the same verb.

Maria and Eve are able to produce novel verb forms based only on information from other verbs or general knowledge of verb conjugation and distributional properties. This can be considered as an indicator for productivity, since morphological information from other verbs may be used to produce novel forms of a verb, but this does not necessarily happen on a verb-by-verb basis.

7.2 Root Infinitives

The aim of the analyses of RIs was to examine the use of non-finite forms in child Greek and investigate the claim that there are two stages in children's speech in regards to the acquisition of finiteness.

The first question to be addressed, from a generative point of view, was whether children use non-finite (bare) verbal forms in main clauses in their early speech production; if yes is this evidence for two stages in the acquisition of early Greek?

This first question explored the production of finite and non-finite non-adult clauses in Maria's and Eve's corpora, building on previous analyses for RIs in child Greek (VVR, 1996/1998; Hyams, 2002, Varlokosta, 2002/2005). The predictions for the first question are that Maria's and Eve's speech develops in two stages: during the first stage, children use a non-finite non-adult verbal form in incorrect contexts while during stage two these non-finite non-adult like forms are reduced and soon disappear.

The second question originates from a usage-based approach and explores if children's early non-adult constructions (RIs) is the result of frequency of some structures from the adult input or simply the result of morphological ambiguity of the verbal system. The predictions to the second question are that frequently used items in the input will be frequently used in the children's speech production and that ambiguous

non-finite verbal forms may result in the production of non-adult construction in children's speech.

The results of the analyses performed are summarised in the following paragraphs.

VVR (1996, 1998) in their study of RI stage in Greek speaking children propose the existence of two stages of their speech development. During Stage 1, children overuse a non-finite non-adult verbal form with the suffix *-i* which corresponds to 3sg person. During Stage 1, 3sg verbs with a different suffix are rarely used. Most of the *-i* forms used during Stage 1, are used incorrectly, i.e. in a non-3sg context and occur mostly with null subjects. VVR (ibid.) also claim that the overuse of the *-i* form occurs mostly with the Perfective stem than with the Imperfective stem. During Stage 2 the *-i* forms are used less frequently than at the first stage and mostly in appropriate context of 3sg.

In relation to VVR's suggested evidence, the following analyses were performed for the study of non-finite non-adult forms in Maria's and Eve's data (discussed in section 5.8, but repeated below for convenience):

An error analysis was completed to check the errors performed in Maria's and Eve's speech that might be related to non-finite forms, e.g. missing modal particles. Also, an analysis of agreement of *-i* form was also conducted in order to identify those *-i* forms used in non-3sg context. The results of the analysis of error rate suggest that Maria's and Eve's overall error rate appears to be at a similar level compared to children's agreement errors in other languages such as German, Italian, Spanish, Catalan etc. (see table 5.5). Such error rate is not as high as suggested to be in VVR's study (see also section 5.5.3 where according to Hyams children's error is very low since they have developed inflection related categories from a very early age).

An analysis of overuse of the *-i* forms (*-i* form 3sg and other 3sg) was performed to test the existence of two stages where the overuse of *-i* forms decreases with time. Also, a stem analysis was performed to check if *-i* forms are used more with the perfective stem rather than the imperfective one. The results of the analysis of overuse of the *-i* form 3sg showed that there is evidence of overuse of the *-i* form 3sg in Maria's and Eve's speech production. Such overuse is slightly more prominent with the perfective form in Maria's speech. In Eve's speech, *-i* forms are used mostly with the imperfective stem. Single-stem 3sg verbs are used in a high rate in both Maria's and Eve's corpora. There is not sufficient evidence in Maria's and Eve's corpora of overuse of the perfective stem in 3rd singular verbs with *-i* suffix.

An analysis of null subjects in *-i* forms was carried out to establish whether *-i* forms occur with null or realised subject. The analysis of null subject in *-i* forms provided no supportive evidence to the claims made by VVR for Stage 1.

In Varlokosta (2002), two phenomena in children's grammar that are suggested to constitute evidence related to the RI stage in Greek are:

- Clitic omission with the use of non-finite forms with perfective aspect with the suffix *-i*
- Reduced use of past tense during the RI stage

To address this issue, an analysis of Past tense and Clitic Object Omission in relation to *-i* forms was performed. The results revealed that past tense is productive from very early in Maria's and Eve's speech and tense distinction is in operation. Past tense is, however, limited to eventive predicates (used with perfective aspect), whereas the use of past tense with imperfective aspect is limited.

The examination of the use of clitics in relation to the RI stage showed that clitic object omission in Maria's and Eve's speech is very low, contrary to Varlokosta's suggestions.

Hyams (2002, 2005), proposed an alternative analysis under the term *bare perfectives* against the participial analysis discussed in VVR. According to Hyams (ibid.), children use a form, called *bare perfective*, which in adult language, is expressed by the *na/ta* clause, i.e. the subjunctive mood, except that it lacks the modal particle. Hyams also proposes that perfective *-i* forms (3sg) are also used with incorrect agreement more often than the other persons. Hyams claims that agreement is productive and the two stages proposed in VVR are not confirmed empirically.

Building on Hyams, a further analysis was performed to complement the analyses mentioned above: a modal particles analysis was carried out in order to establish the rate of use of subjunctive and future constructions and the rate of omission of such particles in obligatory contexts. The findings of the analyses of modal particles use revealed that *na/ta* are used in a productive way in both Maria's and Eve's speech. Omission of the modal particles is lower than the predicted levels and it decreases with time.

In Hyams (2007) it has also been suggested that BPs have a modal or irrealis meaning and they are restricted to eventive predicates. A qualitative analysis of the temporal interpretation of BPs in Maria's and Eve's speech showed that all BPs are eventive, and all except one are telic. In terms of telicity, the prediction that all verbs when combined with perfective aspect become telic was not borne out from this analysis, since all verbs are eventive and there are no examples of stative verbs.

A context analysis of the modal interpretation of BPs suggested that 15/20 of all BPs in the corpora appear to have modal/future interpretation; four BPs refer to past

events and one is unclear but with a modal/future connotation. Overall, the findings of the investigation of cases of Bare Perfectives revealed that there are only few cases (20/3353) in Maria's and Eve's speech, and hence, the results are not conclusive.

To summarise, the study of RIs was motivated by the claim that Greek speaking children may go through a RI stage in their early speech productions. Maria and Eve use a non-finite non-adult form, referred to as the Bare Perfective. However, in their speech production there are only few instances of such forms i.e. 20/3353 in total. Those BPs are mostly used with modal meaning to express desire, wish, deontic etc. and all verbs involved are eventive and telic verbs (as suggested in Stephany and Voeikova, 2003; Varlokosta, 2002/2005; Hoekstra and Hyams, 1998).

In terms of syntactic analysis of BPs, the participial analysis proposed by VVR (1998) was not borne out in this study. Most BPs in Maria's and Eve's data are used with a modal meaning; this finding rather advocates for the modal analysis proposed by Hyams (2002).

Maria's and Eve's non-finite non-adult forms occur only for a very short period at around the age of 2; therefore by the age of 2;5 both Maria and Eve outgrow the use of these forms. This corresponds to Stage II (MLU: 2 – 2.5) in Brown's classification. Maria's and Eve's speech present a very low rate of incorrect use of verbal morphology and consequently there is no other evidence to suggest the presence of an optional infinitive stage. Therefore, there is no firm evidence for the presence of two stages in the development of their speech production, but this may be the case because the data analysed in this thesis are not from Stage I.

The results also suggest that Maria's and Eve's speech is at a more advanced stage of development compared to the data analysed in previous research (cf. VVR, 1998; Hyams 2002). Maria's speech appears to be in a more advanced stage compared

to Eve's. Eve's data provide us with a developmental representation capturing a wider range of stages, i.e. Stages II, III and V.

The second question, originated from a usage-based point of view, inspired the frequency analysis of the adult input and its effect to the children's data.

In Freudenthal, et al., (2005), MOSAIC was used to simulate children's speech. The prediction of the procedure is that RIs observed in children's production in various languages, is the result of children omitting sentence-internal elements. This is the direct result of frequency effects, i.e. children do not process low frequency items, usually sentence internally.

The analysis provided counterevidence for this prediction. The examination of the adult production from the recordings constitutes clearly only a small segment of what children hear when interacting with adults. These findings by no means can account for the adult input as a whole or depict a realistic picture of the role of input in language acquisition.

The study of Kupisch and Rinke (2007) examined adults' and children's data in various languages, such as Italian, Brazilian Portuguese, French, German, and English. Their results suggest that ambiguity in the verbal morphology of a given language may cause delays in the production of finite verb. Thus the morphological properties of a language may be related to the rate of RIs produced by children learning this language.

A morphological study of ambiguous verbal forms in MG suggests that one ambiguous form may be the cause of the short RI stage encountered in Maria's and Eve's speech being considered as a learning problem. The assessment of the morphological complexity of the MG verbal system, following the study by Kupisch and Rinke's (2007), suggested that the ambiguity between two verbal forms as discussed in Chapter 5, may be the cause of the presence of non-finite non-adult forms

in Maria's and Eve's speech. However, there is no strong evidence for such a hypothesis because of the very few instances of these forms in Maria's and Eve's data. The general developmental representation we now have through Maria's and Eve's data analysis may also suggest that the rich morphological marking available in MG facilitates the process of acquisition and the early grammatical development of Greek speaking children (see also Xanthos et al., 2009).

7.3 Number and Person marking

The study of person and number morphology was inspired by previous research on inflectional morphology, and especially studies examining the emergence and use of person and number markings by very young children. Maria's and Eve's spontaneous speech data offered a unique opportunity to examine the emergence and productivity of the inflectional paradigm. The study of number and person marking addressed three questions:

The first question examined the first use, time of acquisition of person and number morphology distinctions, and the productivity of these verbal markings. Generative theories (cf. VEPS, VEKI) predict that children acquire inflectional morphology very early and they use the inflectional paradigm productively. Research in MG, however, suggested that there are ordering distinctions, i.e. 3rd person emerges/is acquired before 1st and 2nd and singular emerges/is acquired before plural and that children do not use person and number marking productively from the onset.

Usage-based theories predict that children's acquisition of inflectional morphology is characterised by three stages, namely, the pre-, the proto and the morphology proper stage, and therefore, acquisition is not instant, but subject to a

gradual development. Also, children do not use person and number morphology productively from the onset (cf. Verb Island Hypothesis).

The second question looked at children's substitutions of person and number marking, e.g. 1st person with 3rd and plural with singular. Generative theories predict that children do not substitute inflectional markings, although studies in MG within the maturation approach, predict that children substitute 1st and 2nd person with 3rd and plural with singular.

The third question, from a usage-based point of view explored the effect of frequency, i.e. are more frequently used verb forms of person/number in adults more frequent in children too? The predictions for question three are that children's frequency of use of verbs is closely related to the frequency adults use these verbs.

The results of the analyses performed are summarised below.

An error analysis performed to examine the accuracy of use of person and number morphology in Maria's and Eve's speech showed that the overall rate of agreement error is very low in both Maria's and Eve's speech, i.e. 1.9% and 0.7% respectively. At Maria's first and Eve's second recordings, however, the agreement error rate is much higher, namely, 7.9% and 7.3% respectively. 3rd singular person provides the context for most errors both in Maria's and Eve's speech. Error rates decrease with age. Both singular and plural morphology are present and produced in Maria's and Eve's speech from the onset of the data collection. Plural, however, appears to have limited contexts in both corpora compared to singular.

A second analysis of agreement errors was performed to examine the context of these errors, namely, substitution of one person and number verbal suffix for another. The analysis revealed that both Maria's and Eve's agreement error rates are low and accuracy of agreement morphology increases with age. Maria uses 3sg instead of 2sg

person while Eve uses 3sg in plural contexts. Both Maria and Eve substitute 3sg with 1sg and vice versa.

According to Wexler (1996), children at around 18 months of age, set correctly the basic parameters of the language they learn. This is referred to as the Very Early Parameter Setting (VEPS). Wexler (1998) also postulated the Very Early Knowledge of Inflection (VEKI) according to which children know the inflectional components in the language they learn at a very early stage of language acquisition (from around 18 months of age). It has also been suggested that agreement morphemes are known to very young children (Wexler, 1990, 1994). Hoekstra and Hyams (1995) also claimed that children acquire the specifics of inflection of the target language at a strikingly early age. Hoekstra and Hyams (*ibid.*) refer to this finding as Early Morphosyntactic Convergence (EMC).

The findings of the data analysis suggest that inflectional morphology is used most of the time correctly and the rates of error are very low. These results conform to claims made for children learning other languages and also provide support to the VEPS, VEKI, and EMC hypotheses.

Pizzuto and Caselli (1994), proposed two conditions for productivity (repeated here for convenience):

- a. The same verb root appears in at least two distinct inflected forms and;
- b. The same inflection is used with at least two different verbs.

(Pizzuto and Caselli, 1992)

According to these conditions, an inflectional form is considered to be productive when either or both conditions hold. Guasti's (1994) productivity criteria are very similar. Guasti examined Martina's speech for different verbs with the same

agreement morphemes and the use of the three singular agreement markers with the same verbal root.

Based on these productivity criteria, a qualitative analysis was conducted to address the issue of productivity of the use of person and number suffixes. The findings of this analysis suggested that Maria and Eve have started to build a verbal paradigm analysing the verb into its root and its inflectional affix from the onset of the data collection and they both use the inflectional paradigm productively.

Katis (1984), from the results of her longitudinal and cross-sectional study, suggested the order of emergence of the person and number categories to be singular before plural and 3rd person before 1st and 2nd. Tsimpli (1992) claimed that during the prefunctional stage, children use mostly the 3rd singular person for most of the verbs they produce.

In the context analysis, the high rate of accuracy suggests that inflectional morphology is used productively from a very young age. Children at the earliest age appear to commit some errors mainly when the use 3rd singular is involved. All persons and number markings appear to emerge from the onset of the data collection and they are used productively thereafter. There is, however, more singular context than plural.

The findings of this analysis are different to Katis' and Tsimpli's findings. The patterns of emergence may suggest that Maria and Eve are at a later stage of grammatical development compared to the children analysed in their studies.

Dressler and Karpf (1995), Dressler (2003), Bittner et al. (2003), amongst others supporting the usage-based approach, proposed that children's morphological development is divided into three stages, namely, pre-morphology, proto-morphology, and morphology proper. During the pre-morphological stage, children learn by rote while they only start constructing morphological patterns by analogy at the proto-

morphological stage. Adult like morphology is only achieved at the morphology proper stage where children create complex morphological constructions (Dressler, 2003; Bittner et al., 2003).

Stephany (1985) suggested that the lack of contrasting inflectional forms of the same verb in children's early speech defines the pre-inflectional stage. The results of the study of Christofidou and Stephany (2003) of two databases (Christos from 1;7.11 to 2;1.27 and Mairi from 1;9.17 to 1;9.29) suggested that the onset of the proto-morphological stage for inflectional development, is placed at around the age of 1;11. According to Christofidou and Stephany (*ibid.*), Mairi's speech can be already positioned in the initial phases of the inflectional stage whereas Christos' verbal and inflectional development is at the pre-morphological stage (1;7 – 1;10).

Based on Stephany (1985), Dressler and Karpf (1995), Dressler (2003), Bittner et al. (2003), Christofidou and Stephany (2003), from the analysis of their speech, Maria and Eve appear to be at the morphology proper stage, i.e. the adult-like stage of morphological use and knowledge, from the onset of the data collection. This finding is different from what is proposed for Christos' speech in Christofidou and Stephany (*ibid.*), but similar to what they proposed to be Mairi's developmental stage. The onset of morphological stages of development can be assumed to be individual to each child's analysed speech production.

According to the Stem Parameter proposed by Hyams (2008), children learning a rich in inflection language acquire its inflectional system early and with fewer errors than children learning a poorer in inflection language. Xanthos et al. (2009) argued that variation in child-directed speech has a positive effect on children's early development of inflection. The results of Xanthos' et al. (*ibid.*) study showed that children exposed to

morphologically rich languages acquire morphology at a greater speed compared to children learning morphologically less rich languages Xanthos et al. (ibid.).

The morphological richness of MG, as claimed in Xanthos et al. (2009) may provide children with a fertile ground for easier and quicker acquisition of inflectional morphology (cf. also Hyams, 2008).

Usage-based accounts maintained that frequency of occurrence (adult input) plays a vital role in language acquisition, the assumption being that children select highly frequent items from what they hear and then construct patterns as a result of this natural selection from the input (Bybee, 1991, 1995, 2007; Bybee and Hopper, 2001; Tomasello, 2000, 2003, 2006, amongst others).

Based on the above, and to answer the third question mentioned above regarding frequency patterns, a frequency analysis was performed to examine the use of singular morphology in Maria's and Eve's data and thereafter compare this to the adults' speech in each dataset. The results showed that Maria's and Eve's use of singular morphology is very similar for each person with the 3sg being the most frequently used person. Maria's and Eve's use of singular morphology, however, is not similar to the adult use of singular, i.e. adults use more frequently the 2sg and less frequently the 1sg than Maria and Eve do. Therefore frequency effects are not observed in the use of singular morphology in the corpus. Maria's and Eve's distribution of use of singular verbs are very similar whereas there is not similarity with the adult language. This finding is interesting in relation to studies on frequency discussed above.

7.4 Limitations of the study and future research

Research conducted using data, experimental or longitudinal, has strengths and limitations which should always be taken into consideration.

The present study has used a longitudinal design because its main purpose was to examine the language development of children's spontaneous speech with emphasis on the growth of the verbal domain. Apart from the ability to capture the grammatical development of children's spontaneous speech, longitudinal studies, also allow for quantitative analysis. The advantage of quantitative analysis is that the researcher can obtain information about constructions from different domains, and compare different phenomena at the same time and evaluate the effects they may have to each other.

Longitudinal data, however, have some limitations too. It is very difficult to elicit and investigate phenomena of grammar that do not appear in every day spontaneous speech production. Longitudinal data may not provide any data if the child is unwilling to speak during a recording or when children do speak, the data may be impossible to interpret, especially when children are very young.

Following the above, further research should focus on the exploration of infrequent phenomena of the verbal domain by investigating language production through experimental methodology. For example, testing the children's production and comprehension of past tense with elicitation tasks would strengthen the evidence provided in this thesis even more or the use of cross-sectional data can test the effects of frequency on person and number morphology by investigating language production and comprehension. Additionally, elicitation and cross-sectional tests would shed more light into the production of non-finite non-adult verbal forms in children's early speech.

The present research raises questions about the similarities of the two main theories of language acquisition. Specific elements of each theory have often a parallel focus. An interesting new line of research would be to extend the examination of elements from both frameworks and to attempt to provide integrated accounts and to

attempt to bridge the gap between the two theories, by bringing together those similar elements that matter most for the process of language acquisition.

Most studies of language acquisition examine a small amount of data. The data analysed in this thesis, although used for the first time, are also of limited scope in terms of providing firm conclusions. Another interesting extension of the work presented here would be to collate and analyse more data in MG.

7.5 Conclusion

Overall this thesis has provided evidence that children use verb constructions from a very early age and mostly productively. Children's errors may be the result of some morphological ambiguity; however, largely error rates are very low.

Maria's and Eve's grammatical knowledge is at an advanced stage of language development. The time of acquisition and emergence of specific elements of the verbal domain could vary across children and also in relation to input or usage of specific grammatical components.

The study of the acquisition of past tense showed that the use of sigmatic vs. non-sigmatic morphology varies over time: early on, children use more sigmatic than non-sigmatic forms. The use of non-sigmatic forms increases and in the final recordings the use of sigmatic past tense increases again. The overall results suggest that non-sigmatic forms (irregular formation) are used less often than sigmatic ones (regular formation), especially at the earliest production. In relation to Mastropavlou's classification of salience, the findings of this analysis provide no support of the claim that [+salient] items are easier to acquire than [-salient] ones. The findings of frequency analysis in Maria's and Eve's data revealed that children use frequently those verbs that are frequently used in the adult language (from the Hellenic National Corpus). Children

also produce the verbs they know and used before by expanding on these constructions, as suggested by the trace-back procedure.

Analyses of the use of non-finite non-adult constructions (i.e. RIs) have shown that Maria and Eve use few non-finite non-adult verbal forms for a short period early in their language development; these forms are referred to as the Bare Perfective and their use disappears overtime. Those BPs are mostly used with modal meaning to express desire, wish, deontic meaning etc. and all verbs involved are eventive and telic verbs. Therefore, this finding provides support for the modal analysis proposed by Hyams (2002). The analysis of adult speech in the data collection demonstrated that input does not play a role in the production of RIs in children's speech.

Finally, the investigation of person and number morphology showed that the overall rate of agreement error is low (high at the beginning of the data collection but decreases overtime). 3rd singular person provides the context for most errors both in Maria's and Eve's speech. All persons and number markings appear to emerge from the onset of the data collection and they are used productively thereafter. There is, however, more singular context than plural. A frequency analysis performed to examine the use of singular morphology in Maria's and Eve's data, and thereafter, compare this to the adults' speech in each dataset, showed that Maria's and Eve's use of singular morphology is very similar for each person with the 3sg being the most used person. Maria's and Eve's use of singular morphology, however, is not similar to the adult use of singular.

The results show that Greek speaking children use, mostly accurately, tense, inflection and subject-verb agreement from a very early age of their language development. These findings provide new evidence in the field of language acquisition and support theories of early language acquisition of the verbal domain.

From a theoretical point of view, this thesis examined the two leading approaches of language acquisition. A usage-based approach posits emphasis on the input and the learning mechanism while minimises the role of innate endowment of the child. A generative approach, on the other hand, attributes the success of language acquisition to the innate knowledge of natural language (i.e. UG) while the learning mechanism is reduced in importance. Although this study introduced a new methodology where both approaches were considered and analysed, the former approach did not offer sufficient empirical evidence to assume that the sole use of the learning mechanism alongside the input can provide an adequate language acquisition theory.

The usage-based assumption that children extract statistical information from the input, such as frequency, was not always borne out of the analyses provided in this study, as much of the complexities in past tense, RIs and person and number agreement were not empirically accounted for by this approach. Child and adult languages present significant differences in statistical distribution and therefore the usage-based approach has not tackled the problem of language acquisition in an empirical context. At the same time, the early competence of children's language, as analysed through the data of Maria's and Eve's speech production, provides empirical evidence for the generative approach where a richly endowed innate state can be assumed.

APPENDICES

Appendix A: Are verbs islands? – An analysis and evidence against the Verb Island Hypothesis

Tomasello (1992) used a diary in order to document the development of his daughter's Travis verbal domain starting at 1;3.20 until the age of 1;8.08. The analysis of Tomasello's study used trace-back⁵⁰ procedures designed to trace antecedents of three word sentences back to previous production. Using corpora of speech production, trace-back methods identify multiword utterances and examine how these utterances are related to what the children has said before (Dabrowska and Lieven 2005; Lieven 2006; Lieven et al. 2003).

The results of Tomasello's study suggested that approximately 92% of the child's first three-or-more-word sentences involved only a single simple change from previous sentences with the same verb whereas only 4% could not be traced back to an antecedent (Tomasello, 1992). Tomasello's results are inconsistent with findings from other studies mentioned earlier (Ninio, 2003; Abbot-Smith and Behrens, 2006; McClure and Pine 2002, cited in Ninio, 2003, p. 6), in that, these studies do not support the insularity element of the Verb Island hypothesis. These studies, demonstrated that children's knowledge of constructions and specific patterns facilitate their learning of new verbs, and thus, the connections by analogy and transfer of learning is a vital process of their language development.

Ninio (2003) re-examined the child's data used in Tomasello's study in order to identify those sentences that the child produced for the first time, i.e. without a previous antecedent. Ninio's (ibid.) findings estimated that 40% of the child's new patterns have

⁵⁰ For a detailed discussion of trace-back procedures see chapter 4.

no antecedent in her previous production. This is ten times higher than Tomasello's reports (Ninio, 2003).

Building on Tomasello's methodology, a trace-back procedure is used in the present study to examine one simple construction and its development in Maria's and Eve's speech production throughout different recordings (and different ages).

The data used in this study were not originally designed for a trace-back analysis, so there are certain limitations on the way the data can be used for this procedure. There is, however, scope on performing this procedure since there are no prior usage-based studies and analyses for MG.

Given that many studies providing contra-evidence on the Verb Island Hypothesis examined a sole construction, and due to time restrictions, I have selected a sole construction to trace-back, namely, the subjunctive *na*-construction. The subjunctive *na*-construction was selected because is a widely attested construction in the corpus and because it provides the context for a variety of verbs to be constructed with.

The procedure designed is much simpler than the trace-back system used in previous usage-based studies. It traces the use of subjunctive constructions across inter-age recordings. Matching and deriving operations are not used.

The analysis simply aims to identify the use of *na*-constructions in various points in Maria's and Eve's speech production with the purpose to discuss the variety of verbs, arguments and pattern used with the same construction providing support for the interconnectedness and transfer of knowledge from one verb to another.

The subjunctive in MG is formed with the use of the modal particle *na*⁵¹ preceding the verb. The subjunctive can be used in a main clause or embedded to express wish, desire, request, order, permission etc. from the speaker's point of view

⁵¹ Sometimes also the particle *as* is used to express exhortation, admonition, or consent.

(Holton et al., 1997; Triantafillidis, 2002; Klairis and Babiniotis, 2004). The examples below present subjunctive constructions used both in main and embedded contexts:

1. na to k(l)isume.
toSUBJ it **switch-off/1pl**
Let's switch it off

(Maria 2;0)

2. na to valume pali, ne edo.
toSUBJ it **put/1pl** again, yes here
Let's put it once again; yes here.

(Maria 2;0)

3. thelo ato na to tipao.
want/1sg this toSUBJ it **perforate/1sg**
I want to perforate (it) this

(Maria 2;2)

4. ses na to petatso ato, ato?
Want/2sg toSUBJ it **throw-away/1sg** this, this?
Do you want me to throw it away, this?

(Maria 2;2)

5. tora les esi: selis na to filas to funtuli?

Now **say/2sg** you: **want/2sg** toSUBJ it **kiss/2sg** the funtuli⁵²?

Now you say: do you want to kiss it, funtuli?

(Maria 2;2)

6. afu thes na to valis ato gia na kimithune.
after **want/2sg** toSUBJ it **put/2sg** this so toSUBJ **sleep/3pl**

Since you want to put this so they can sleep.

(Maria 2;2)

7. ke ato tora pu na to valo?
and this now, where toSUBJ it **put/1sg**?

So, where shall I put this now?

(Maria 2;2)

8. thelo na, thelo na to patiso.
Want/1sg toSUBJ, **want/1sg** toSUBJ it **push/1sg**

I want to, I want to push it.

(Maria 2;3)

9. thelo na to patisis, ela patise to.
want/1sg toSUBJ it **push/2sg**, **come3sg** **push/2sgIMP** it.

I want you to push it; come on push it.

(Maria 2;3)

⁵² A doll's name.

10. thelo na to kliso na to paris.
Want/1sg toSUBJ it **close/1sg** toSUBJ it **take/2sg**
I want to close this so you can take it

(Maria 2;3)

11. momo selo na to siakso
mummy **want/1sg** toSUBJ it **fix/1sg**
mummy, I want to fix it.

(Maria 2;3)

12. (th)elo na to kuposu.
want/1sg toSUBJ it **button/1sg**
I want to button it

(Eve 1;10)

13. de boro na to kuposu.
not **can/1sg** toSUBJ it **button/1sg**
I cannot button it

(Eve 1;10)

14. pame na to v(r)ume
go/1pl toSUBJ it **find/1pl**
Let's go find it.

(Eve 1;11)

15. ela na to dume.

Come/3sg toSUBJ it **see/1pl**

Come on, let's see it

(Eve 2;0)

16. theli na to ftjatsi to tokinito tu baba.

Want/2sg toSUBJ it **fix/3sg** the car ofGEN daddy

He/she wants to fix daddy's car

(Eve 2;0)

17. pame sto zepelin na to dume.

go/1pl to-the zeppelin toSUBJ it **see/1pl**

Let's go and see the zeppelin.

(Eve 2;0)

18. na to ritso edo mesa?

toSUBJ it **drop/1sg** here inside?

Shall I drop this inside here?

(Eve 2;2)

19. na to skupiso me afta?

toSUBJ it **wipe/1sg** with these?

Shall I wipe it with these?

(Eve 2;2)

20. ego na to ftjatso.

I toSUBJ it **fix/1sg**

I shall fix it.

(Eve 2;2)

21. thelo na to anikso.

want/1sg toSUBJ it **open/1sg**

I want to open it

(Eve 2;2)

22. de thelo na to po.

not **want/1sg** toSUBJ it **say/1sg**

I don't want to say it.

(Eve 2;2)

23. na to ferume.

toSUBJ it **bring/1pl**

We shall bring it

(Eve 2;2)

The examples above, from Maria's and Eve's speech present the use of subjunctive constructions with the particle *na* in main and embedded contexts.

Subjunctive constructions are used in main and embedded contexts, with different verbs and with different person and number inflectional markers, and a surplus of other words such as adverbs, adjectives etc. The richness and variety of the patterns suggests that Maria and Eve are able to use the same construction they learned at an early age, i.e. 2;0 for Maria and 1;10 for Eve, with a variety of different verbs and in a variety of patterns. Therefore, Maria and Eve can transfer by analogy their knowledge of one pattern used with one verb to other, more enhanced patterns, with more arguments and used with different verbs. The productivity of this sole pattern across Maria's and Eve's corpora challenges the Verb Island Hypothesis, demonstrating that verbs are not isolated constructions with their own internal syntax learned on an item-by-item basis. Verbs' combinatorial and syntactic information is available to children, at an early stage, to transfer and use with other verbs, and therefore the verbal system is interconnected. This corresponds to the second stage of development proposed in Verb Island Hypothesis by Tomasello (1992).

The findings of a simple trace-back analysis presented above suggest that Maria and Eve use analogical processes to transfer what they learned for one verbal construction to another. The examples presented provide evidence for Maria and Eve being in a more advanced stage of their speech development.

REFERENCES

- Abbot-Smith, K., and Behrens, H. (2006). How known constructions influence the acquisition of other constructions: The German passive and future constructions. . *Cognitive Science*, 30(6), 995-1026.
- Adger, D. (2003). *Core syntax: a minimalist approach*. Oxford: Oxford University Press.
- Agouraki, Y. (1991). A Modern Greek complementiser and its significance for UG. *UCLWPL*, 3.
- Alexaki, C. (2003). *Erstspracherwerb des Neugriechischen: Eine Fallstudie zum Erwerb von Verbformen* Unpublished PhD Thesis, University of Salzburg.
- Atkinson, M. (1992). *Children's syntax: An introduction to principles and parameters theory* (Vol. 5). Oxford: Blackwell.
- Bel, A. (2001). *The projection of aspect: A key in the acquisition of finiteness?* . Paper presented at the Proceedings of the 8th Conference of the International Association for the Study of Child Language.
- Bittner, D., Dressler, W. U., and Kilani-Schoch, M. (2000). First verbs: On the way to Mini-paradigms. *ZAS Papers in Linguistics*, 18.
- Bittner, D., Dressler, W. U., and Kilani-Schoch, M. (2003). *Development of verb inflection in first language acquisition : a cross-linguistic perspective*. Berlin: Mouton de Gruyter.
- Blom, E. (2003). *From root infinitive to finite sentence*. Unpublished Ph.D., Utrecht Institute of Linguistics OTS.

- Blom, E., and Wijnen, F. (2006). *Development need not be embarrassing: The demise of the root Infinitive and related changes in Dutch child language*. Unpublished manuscript, University of Amsterdam/Utrecht University.
- Bloom, L. (1970). *Language Development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.
- Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax*. The Hague, The Netherlands: Mouton.
- Borer, H. (1984). *Parametric syntax : case studies in Semitic and Romance languages*. Dordrecht: Foris.
- Borer, H., and Wexler, K. (1987). The maturation of syntax. In T. Roeper. and E. Williams (Eds.), *Parameter-setting and language acquisition*. Dordrecht, The Netherlands: Reidel.
- Boser, K., Lust, B., Santelmann, L., and Whitman, J. (1992). *The syntax of CP and V-2 in early child German (ECG) - the strong continuity hypothesis*. Paper presented at the Proceedings of the North Eastern Linguistic Society 22, University of Massachusetts, Amherst.
- Brown, R. W. (1973). *A first language: The early stages*. Cambridge, Mass.: Harvard University Press.
- Bybee, J. (1991). Natural morphology: the organization of paradigms and language acquisition. In T. H. a. C. Ferguson (Ed.), *Crosscurrents in Second Language Acquisition and Linguistic Theories*. Amsterdam: Benjamins.
- Bybee, J. (1995). The Semantic Development of Past Tense Modals in English. In J. Bybee and S. Fleischman (Eds.), *Modality in Grammar and Discourse* (pp. 503-517). Amsterdam: Benjamins.

- Bybee, J. L. (2007). *Frequency of use and the organization of language*. New York ; Oxford: Oxford University Press.
- Bybee, J. L., and Hopper, P. J. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam ; [Great Britain]: J. Benjamins.
- Cameron-Faulkner, T., Lieven, E., and Tomasello, M. (2003). A Construction Based Analysis of Child Directed Speech. *Cognitive Science*, 27, 843-873.
- Campos, H. (1989). Modern Greek and CP Transparency. *Linguistic Inquiry*, 20, 321-327.
- Carnie, A. (2002). *Syntax : a generative introduction*. Oxford: Blackwell.
- Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development*, 39, 433-448.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Mass.: M. I. T. Press.
- Chomsky, N. (1976). *Reflections on language*. London: Temple Smith.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1984). *Modular approaches to the study of the mind*: San Diego State University Press.
- Chomsky, N. (1986). *Knowledge of language : its nature, origins, and use*. New York: Praeger.
- Chomsky, N. (1988). *Language and Problems of Knowledge. The Managua Lectures*. Cambridge, Mass: MIT Press.
- Chomsky, N. (1992). *A minimalist program for linguistic theory. MIT occasional papers in Linguistics 1*. Cambridge, Mass: MIT Press.
- Chomsky, N. (1995). *The minimalist program*. Cambridge Mass: MIT press.
- Chomsky, N. (1998). Minimalist inquiries: The framework. *MIT Occasional Papers in Linguistics*, 15.

- Chomsky, N. (1999). *Derivation by Phase*. Cambridge, Mass: MIT Press.
- Chomsky, N., and Collins, C. E. I. (2001). *Beyond explanatory adequacy*. Cambridge, Mass.: MITWPL.
- Chomsky, N., and Lasnik., H. (1993). *The theory of principles and parameters* (Vol. 1): Walter de Gruyter.
- Christofidou, A., and Stephany, U. (2003). *Early Phases in the Development of Greek Verb Inflection*. Berlin/New York: Mouton de Gruyter.
- Clahsen, H. (1991). Constraints on parameter setting: A grammatical analysis of some acquisitional stages in German child language. *Language Acquisition*, 1, 361-391.
- Clahsen, H. (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioural and Brain Sciences*, 22(991-1060).
- Clahsen, H., and Stavrakaki, S. . (2004). *The perfective past-tense Test*. University of Essex and Aristotle University of Thessaloniki.
- Clahsen, H., and M, A. (1998). Syntax and morphology in Williams Syndrome. *Cognition*, 68(3), 167-198.
- Clahsen, H., and Penke, M. (1992). *The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-Corpus*: Dordrecht: Kluwer.
- Comrie, B. (1981). *Language universals and linguistic typology : syntax and morphology*. Oxford: Blackwell.
- Crain, S., and Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, Mass: MIT Press.
- Dąbrowska, E., and Lieven, E. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics*, 16(3), 437-474.

- Delidaki, S., and Varlokosta, S. (2003). Testing the aspect first hypothesis: A preliminary investigation into the comprehension of tense in child Greek. In B. D. N. Gagarina (Eds.), *ZAS Papers in Linguistics* (Vol. 29, pp. 73-84).
- Demuth, K. (1996). The prosodic structure of early words. In J. L. Morgan and K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 171-184). Mahwah, NJ: Lawrence Erlbaum Associates.
- Deprez, V. (1994). Underspecification, functional projections, and parameter setting. In B. Lust, M. Suñer and J. Whitman (Eds.), *Syntactic theory and first language acquisition - crosslinguistic perspectives: Volume 1 - heads, projections and learnability*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dockrell, J., Messer, D., and George, R. (2001). Patterns of naming objects and actions in children with word finding difficulties. *Language and Cognitive Processes*, 16(2/3), 261-286.
- Doukas, T. (1998). The null subject parameter and non-finite forms on native Greek language acquisition. Study of a case. Unpublished BA Dissertation. University of Siena.
- Drachman, G., and Malikouti-Drachman, A. (1994). Stress and Greek Compounding. *Phonologica*, 55-64.
- Dressler, W. U. (2003). *Naturalness and morphological change*. Oxford: Blackwell.
- Dressler, W. U. (2005). *Morphology and its demarcations : selected papers from the 11th Morphology Meeting, Vienna, February 2004*. Amsterdam ; Philadelphia, PA: John Benjamins.
- Dressler, W. U., and Karpf, A. (1995). *The theoretical relevance of pre- and protomorphology in language acquisition*. Dordrecht: Kluwer Academic Publishers.

- Eisenbeiss, S. (2009). Generative Approaches to Language Learning. *Linguistics* 47(2), 273–310.
- Ezeizabarrena, M. J. (2002). *Root infinitives in two pro-drop languages*. Dordrecht: Kluwer.
- Felix, S. W. (1991). Language acquisition as a maturational process. In J. Weissenborn, H. Goodluck and T. Roeper (Eds.), *Theoretical issues in language acquisition: Continuity and change in development*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Felix, S. W., and Schumann, J. H. (1984). Maturational aspects of universal grammar. In A. Davies (Ed.), *Interlanguage* (pp. 133-161). Edinburgh: Edinburgh UP.
- Ferdinand, A. (1995). On the acquisition of finiteness in French. In M. Verrips and F. Wijnen (Eds.), *Papers from the Dutch-German Colloquium on Language Acquisition (Groningen 1-2 September 1994)*. Amsterdam, The Netherlands: Institute for General Linguistics, University of Amsterdam.
- Ferguson, C. A. (1978). Talking to children: A search for universals. In J. Greenberg (Ed.), *Universals of Human Language, vol. 1*. Stanford, CA: Stanford University Press.
- Fernald, A., Taeschner, T., Dunn, J., Papoushek, M., de Boysson-Bardies, B., and Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477-501.
- Fowler, A. E., Gelman, R., and Gleitman, L. R. (1994). The course of language learning in children with Down syndrome. In H. Tager Flusberg (Ed.), *Constraints on language acquisition: Studies of atypical children* (pp. 91-140). Hillsdale, NJ: Erlbaum.

- Freudenthal, D., Pine, J. M., and Gobet, F. (2004). *Simulating the temporal reference of Dutch and English Root Infinitives* Paper presented at the Proceedings of the 26th Annual Meeting of the Cognitive Science Society
- Freudenthal, D., Pine, J. M., and Gobet, F. (2005). *Simulating Optional Infinitive errors in child speech through the omission of sentence-internal elements*. Paper presented at the Proceedings of the 27th Annual Meeting of the Cognitive Science Society.
- Freudenthal, D., Pine, J. M., and Gobet, F. (2006). Modelling the development of children's use of Optional Infinitives in English and Dutch using MOSAIC. *Cognitive Science*, 30, 277-310.
- Freudenthal, D., Pine, J. M., and Gobet, F. (2009). Simulating the referential properties of Dutch, German and English Root Infinitives in MOSAIC. *Language Learning and Development*, 5, 1-29.
- Fromkin, V., Rodman, R., and Hyams, N. M. (2003). *An introduction to language* (7th ed.). Boston, Mass. ; London: Thomson/Heinle.
- Gerken, L. (1994). Young children's representation of prosodic phonology: Evidence from English-speakers' weak syllable productions. *Journal of Memory and Language*, 31(JMemL).
- Giannakidou, A. (2002). *UNTIL, aspect and negation: a novel argument for two untils*. Paper presented at the Semantics and Linguistic Theory (SALT)
- Giannakidou, A. (2003). *A puzzle about until and the Present Perfect*. Paper presented at the Perfect Explorations (Selected papers from the workshop on the Perfect), University of Thessaloniki, Greece, May 2000.
- Giorgi, A., and Pianesi, F. (1997). *Tense and Aspect: from Semantics to Morphosyntax*: New York/Oxford: Oxford University Press.

- Goodluck, H. (1991). *Language acquisition: A linguistic introduction*. Oxford: Blackwell.
- Greenberg, J. (1954). A quantitative approach to the morphological typology of language. In R. F. Spencer (Ed.), *Method and Perspective in Anthropology* (pp. 192-220): Minneapolis: University of Minnesota Press.
- Guasti, M.-T. (1994). Verb syntax in Italian child grammar: Finite and nonfinite verbs. *Language Acquisition: A Journal of Developmental Linguistics*, 3(1), 1-40.
- Guasti, M. T. (2002). *Language acquisition. The growth of grammar*: MIT Press.
- Gueron, J. (2008). On the difference between telicity and perfectivity *Lingua*, 118, 1816–1840.
- Guilfoyle, E., and Noonan, M. (1992). Functional categories and language acquisition. *Canadian Journal of Linguistics Revue Canadienne de Linguistique*, On, Canada (CJL).
- Haan, G. d. (1986). A theory-bound approach to the acquisition of verb placement in Dutch In G. d. H. W. Zonneveld (Ed.), *Formal Parameters of Generative Grammar* (Vol. 3, pp. 15-30): Utrecht: University of Utrecht.
- Haegeman, L. (1996). Root infinitives, clitics and truncated structures. In H. Clahsen (Ed.), *Generative perspectives on language acquisition: Empirical findings, theoretical considerations and crosslinguistic comparisons* (pp. 271-307). Amsterdam: Benjamins.
- Hamann, C., Rizzi, L., and Frauenfelder, U. H. (1996). On the acquisition of subject and object clitics in French. In H. Clahsen (Ed.), *Generative perspectives on language acquisition: Empirical findings, theoretical considerations and crosslinguistic comparisons* (pp. 309-334). Amsterdam: Benjamins.
- Han, C.-h. (2001). Force, negation and imperatives. *The Linguistic Review* 18, 289-325.

- Harley, T. A. (2001). *The psychology of language : from data to theory* (2nd ed. ed.). Hove: Psychology.
- Harris, T., and Wexler, K. (1996). The optional-infinitive stage in child English: Evidence from negation. In H. Clahsen (Ed.), *Generative perspectives on language acquisition: Empirical findings, theoretical considerations and crosslinguistic comparisons* (pp. 1-42). Amsterdam: Benjamins.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569-1579.
- Hellenic National Corpus. (1999-2009). Institute for Language and Speech Processing: Web Version 3.0
- Hickey, T. (1993). Identifying formulas in first language acquisition. *Journal of Child Language*.
- Hoekstra, T., and Hyams, N. (1995). *The syntax and interpretation of dropped categories in child language: a unified account*. Paper presented at the Proceedings of the West Coast Conference on formal Linguistics
- Hoekstra, T., and Hyams, N. (1998). Aspects of root infinitives. *Lingua*, 106, 81-112.
- Holton, D., Mackridge, P., and Philippaki-Warbuton, I. (1997). *Greek : a comprehensive grammar of the modern language*. London: Routledge.
- Holton, D., Mackridge, P., and Philippaki-Warbuton, I. (2004). *Greek : an essential grammar of the modern language*. London: Routledge.
- Howe, C. J. (1980). Mother-child conversation and semantic development. In H. Giles, P. Robinson William and M. Smith Philip (Eds.), *Language: Social psychological perspectives* (pp. 35-40). Oxford: Pergamon.
- Hyams, N. (1986). *Language acquisition and the theory of parameters*. Norwell, MA.: Reidel.

- Hyams, N. (1992). The Genesis of Clausal Structure. In J. M. Meisel (Ed.), *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition* (pp. 371-400). Dordrecht, London: Kluwer.
- Hyams, N. (1996). The underspecification of functional categories in early grammar. In H. Clahsen (Ed.), *Generative perspectives on language acquisition: Empirical findings, theoretical considerations and crosslinguistic comparisons* (pp. 91-127). Amsterdam: Benjamins.
- Hyams, N. (2001). *Now you hear it, now you don't: the nature of optionality in child grammar*. Paper presented at the Proceedings of BUCLD.
- Hyams, N. (2002). Clausal Structure in Child Greek: A reply to Varlokosta, Vainikka and Rohrbacher and a reanalysis. *The Linguistic Review* 19.
- Hyams, N. (2005). Child Non-Finite Clauses and the Mood-Aspect Connection: Evidence from Child Greek. In R. S. P. Kempchinsky (Ed.), *The Syntax, Semantics and Acquisition of Aspect*. Dordrecht: Kluwer Academic Publishers.
- Hyams, N. (2007). Aspectual Effects on Interpretation in Early Grammar. *Language Acquisition*, 14(3), 231–268.
- Hyams, N. (2008). The acquisition of inflection: A parameter-setting approach. *Language Acquisition*, 15(3), 192-209.
- Ingram, D., and Thompson, W. (1996). Early syntactic acquisition in German: Evidence for the modal hypothesis. *Language*, 72, 97-120.
- Jordens, P. (1990). The acquisition of verb placement in Dutch and German. *Linguistics*, 28(6).
- Katis, D. (1984). The acquisition of the modern Greek verb : With special reference to the imperfective past and perfect classes. [electronic resource].

- Klairis, C., and Babinotis, G. (2004). *Grammatiki tis Neas Ellinikis (Grammar of Modern Greek)*. Athens: Ellinika Grammata.
- Kupisch, T., and Rinke, E. (2007). A fresh look at root infinitives from a cross-linguistic perspective. In N. M. Anderssen and M. Westergaard (Ed.), *Tromsø University Working Papers on Language and Linguistics* (Vol. 34 (3)).
- Legate, J., and Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition*, 14(3), 315–344.
- Liceras, J., Bel, A., and Perales, S. (2006). *Living with Optionality: Root Infinitives, Bare Forms and Inflected Forms in Child Null Subject Languages*. Paper presented at the Selected Proceedings of the 9th Hispanic Linguistics Symposium.
- Liceras, J. M., Valenzuela, E., and Díaz, L. (1999). L1/L2 Spanish grammars and the pragmatic deficit hypothesis. *Second Language Research* 15, 161-190.
- Lieven, E. (2006). Producing multiword utterances. In B. K. E. Clark (Ed.), *Constructions in Acquisition* (pp. 83-110): Stanford, CA: CSLI Publications.
- Lieven, E., Behrens, H. Speares, J., and Tomasello, M. . (2003). Early syntactic creativity: a usage-based approach. *Journal of Child Language*, 30, 333-370.
- Lieven, E., Salomo, D., and Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, 20(3), 481-508.
- Lust, B. (2006). *Child language : acquisition and growth*. Cambridge: Cambridge University Press.
- MacWhinney, B. (1978). The acquisition of morphology. *Monographs of the Society for Research in Child Development*, 43(1)2, 1-123.

- MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Marinis, T. (1998). The acquisition of expletive definite articles in Modern Greek. In T. Cambier-Langeveld, A. Lipták and M. Radford (Eds.), *Proceedings of ConSOLE 6, Lisbon, December 1997* (pp. 169-184).
- Marinis, T. (2003). *The Acquisition of the DP in Modern Greek* (Vol. 31). Amsterdam: John Benjamins.
- Mastropavlou, M. (2006). *The Role of Phonological Saliency and Feature Interpretability in the Grammar of Typically Developing and Language Impaired Children* PhD dissertation Aristotle University of Thessaloniki.
- McClelland, J. L., and Patterson, K. (2002). Rules of connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences*, 6, 465-472.
- Meisel, J. M. (1990). INFL-lection: subjects and subject-verb agreement. In M. J. Meisel (Ed.), *Two first languages: Early grammatical development in bilingual children*. Dordrecht: Foris.
- Meisel, J. M. (1995). Parameters in acquisition. In P. Fletcher and B. MacWhinney (Eds.), *The Handbook of Child Language*. Oxford, UK: Basil Blackwell.
- Meisel, J. M., and Ezeizabarrena, M. J. (1996). Subject-Verb and Object-Verb Agreement in Early Basque. In H. Clahsen (Ed.), *Generative Perspectives on Language Acquisition: Empirical Findings, Theoretical Considerations and Crosslinguistic Comparisons* (pp. 201-239). Amsterdam: Benjamins.
- Newport, E. L., Gleitman, H., and Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of motherese. In C. A. Ferguson and C. E. Snow (Eds.), *Talking to children*. New York, NY: Cambridge University Press.

- Ninio, A. (2003). No verb is an island: Negative evidence on the Verb Island hypothesis. *Psychology of Language and Communication*, 7, 3-21.
- Perales, S., Licerias, J., and Bel, A. (2006). Some Remarks on the Realization of the Realis/Irrealis Opposition in Child Language: Towards a Universal Characterization of the Root Infinitive Stage Across Languages *Ottawa Papers in Linguistics - CLO/OPL 34*(Janvier/January 2006), 87-98.
- Perales, S., Spradlin, K. T., and Licerias, J. M. (2004). *More on the issue of optionality: Evidence from finiteness in child L1 and child L2 Spanish*. Paper presented at the Studies in Contrastive Linguistics - Proceedings of the 3rd International Contrastive Linguistics Conference.
- Phillips, C. (1996). Root infinitives are finite. In A. Stringfellow, D. Cahama-Amitay, E. Hughes and A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University Conference on Language Development, Vol. 2* (pp. 588-599). Somerville, MA: Cascadilla Press.
- Pierce, A. E. (1992). *Language acquisition and syntactic theory : A comparative analysis of French and English child grammars*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1990). Language acquisition. In D. N. Osherson and H. Lasnik (Eds.), *Language: An invitation to cognitive science* (Vol. 1). Cambridge, MA: MIT Press.
- Pinker, S. (1991). Rules of language. *Science* 253, 530–535.
- Pinker, S. (1999). *Words and rules: the ingredients of language*. London: Weidenfeld and Nicolson.

- Pinker, S., and Ullman, M. (2002). The past and future of the past tense. *Trends in Cognitive Science*, 6, 456-463.
- Pizzuto, E., and Caselli, M. C. (1992). The acquisition of Italian morphology: Implications for models of language development. *Journal of Child Language*, 19(3), 491-557.
- Platzack, C. (1992). Functional Categories and Early Swedish. In J. M. Meisel (Ed.), *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition*. (pp. 63-82). Dordrecht; London: Kluwer.
- Poepfel, D., and Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language: Journal of the Linguistic Society of America*.
- Pye, C. (1986). Quiche Mayan speech to children. *Journal of Child Language*, 13, 85-100.
- Radford, A. (1988). *Transformational grammar : a first course*. Cambridge Cambridgeshire ; New York: Cambridge University Press.
- Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Oxford, UK: Basil Blackwell.
- Radford, A. (1995). Phrase structure and functional categories. In P. Fletcher and B. MacWhinney (Eds.), *The Handbook of Child Language*. Oxford, UK: Basil Blackwell.
- Ralli, A. (1988). *Eléments de la morphologie du Grec Moderne: La structure du verbe.*, Université de Montréal, Montreal, Canada.
- Ralli, A. (2005). *Morfologia (Morphology)*. Athens: Patakis.
- Rizzi, L. (1993). Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition: A Journal of Developmental Linguistics*, Nj(LAcq).

- Rizzi, L. (1994). Early null subjects and root null subjects. In B. Lust, G. Harmon and J. Kornfilt (Eds.), *Syntactic theory and first language acquisition - Crosslinguistic perspectives: Volume 2 - Binding, dependencies, and learnability*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rizzi, L. (1997). The fine structure of the Left Periphery. In L. Haegeman (Ed.), *Elements of Grammar. Handbook in Generative Syntax* (pp. 281-337). Dordrecht: Kluwer.
- Roussou, A. (2000). On the left periphery: modal particles and complementisers. *Journal of Greek Linguistics, 1*, 65-94.
- Salustri, M., and Hyams, N. (2003). *Is There an Analogue to the RI Stage in the Null Subject Languages*. Paper presented at the Proceedings of the BUCLD
- Salustri, M., and Hyams, N. (2006). Looking for the Universal Core of the RI Stage. In V. Torrens, Escobar, L. (Ed.), *The Acquisition of Syntax in Romance Languages* Amsterdam: John Benjamins.
- Schutze, C. T., and Wexler, K. (1996). Subject case licensing and English root infinitives. In A. Stringfellow, D. Cahama-Amitay, E. Hughes and A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University Conference on Language Development, Vol. 2* (pp. 670-681). Somerville, MA: Cascadilla Press.
- Smith, C. S. (1991). *The parameter of aspect*. Dordrecht ; Boston: Kluwer Academic Publishers.
- Snow, C. E. (1994). Beginning from baby talk: 20 years of research on input and interaction. In C. Gallaway and B. J. Richards (Eds.), *Input and interaction in language acquisition* (pp. 3-12). New York, NY: Cambridge University Press.

- Stathopoulou, N. (2009). *Linguistics profile of Greek individuals with Down Syndrome: evidence from syntactic and morphological phenomena*. Unpublished PhD Thesis, University of Essex.
- Stathopoulou, N., and Clahsen, H. (2010). The perfective past tense in Greek adolescents with Down Syndrome. *Clinical Linguistics and Phonetics* 24, 870-882.
- Stavrakaki, S., and Clahsen, H. (2009). The perfective past tense in Greek child language. *Journal of Child Language*, 36, 113-142.
- Stephany, U. (1981). Verbal grammar in modern Greek early child language. In P. S. Dale and D. Ingram (Eds.), *Child Language - An International Perspective*. Baltimore, Maryland: University Park Press.
- Stephany, U. (1985). *Aspekt, Tempus und Modalität: Zur Entwicklung der Verbalgrammatik in der neugriechischen Kindersprache. [Aspect, tense, and modality: The development of grammar in young Greek children.]*. Tübingen, Germany: Gunther Narr.
- Stephany, U. (1986). Modality. In P. Fletcher and M. Garman (Eds.), *Language Acquisition: Studies in First Language Development* (pp. 375-400). Cambridge: Cambridge UP.
- Stephany, U. (1995). The acquisition of Greek. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition. Vol. 4*. Hillsdale, NJ: Lawrence Erlbaum.
- Stephany, U., and Voeikova, M. D. (2003). On the Early Development of Aspect in Greek and Russian Child Language, a Comparative Analysis. *ZAS Papers in Linguistics* 29, 211 - 224.

- Stephany, U., and Voeikova, M. D. (Eds.). (2009). *Development of Nominal Inflection in First Language Acquisition. A Cross-Linguistic Perspective*. Berlin, New York Mouton de Gruyter.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Tomasello, M. (2000). Do young children have adult syntactic competence? . *Cognition* 74, 209-253.
- Tomasello, M. (2003). *Constructing a language. A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Tomasello, M. (2006). Acquiring linguistic constructions. In D. K. R. Siegler (Ed.), *Handbook of Child Psychology*. New York: Wiley.
- Torrens, V. (1995). *La adquisición de la sintaxis en catalán y castellano: La categoría funcional de flexión*. Unpublished PHD, University of Barcelona, Barcelona, Spain.
- Triantafyllidis, M. (2002). *Neoelliniki Grammatiki (Grammar of Modern Greek)*. Thessaloniki: Modern Greek Studies Institute: Aristotle University of Thessaloniki.
- Tsimpli, I.-M. (1992). Functional categories and maturation : the prefunctional stage of language acquisition. [electronic resource].
- Tsimpli, I. M. (1990). The clause structure and word order of Modern Greek. *UCL working papers in Linguistics* 2, 226-255.
- Tsimpli, I. M. (1996). *The prefunctional stage of first language acquisition: A crosslinguistic study*. New York: Garland.

- Tsimpli, I. M. (2005). Peripheral Positions in Early Greek. In M. S. a. A. Terzi (Ed.), *Advances in Greek Generative Syntax - Linguistics Today* (Vol. 76, pp. 178-216): John Benjamins.
- Tsopanakis, A. (1994). *Neoelliniki Grammatiki (Modern Greek Grammar)*. Thessaloniki, Hellas: Kiriakidi A.E.
- Tsoulas, G. (1994). Remarks on the Structure and Interpretation of Na-clauses. *Studies in Greek Linguistics, XIV*.
- Valian, V. (1999). Rethinking learning: comments on Rethinking innateness. *Journal of Child Language, 26*, 248-253.
- Valian, V. (2009). Innateness and learnability. In E. Bavin (Ed.), *Handbook of Child Language* (pp. 15-34). Cambridge: Cambridge University Press.
- Van Gelderen, V., and Meulen, I. v. d. (1998). Root Infinitives in Russian: Evidence from Acquisition. Leiden University.
- Varlokosta, S. (2002). On the semantic properties of non-finite clauses in early child Greek. *Reading Working Papers in Linguistics, 6*, 133-159.
- Varlokosta, S. (2005). Eventivity, modality and temporal reference in child Greek. In M. Stavrou and A. Terzi (Eds.), *Advances in Greek Generative Grammar*. Amsterdam/Philadelphia: John Benjamins Publishers.
- Varlokosta, S., Vainikka, A., and Rohrbacher. (1998). Functional projections, and 'root infinitives' in early child Greek. *The Linguistic Review, 15*, 187-207.
- Varlokosta, S., Vainikka, A., and Rohrbacher, B. (1996). Root infinitives without infinitives. In A. Stringfellow, D. Cahama-Amitay, E. Hughes and A. Zukowski (Eds.), *Proceedings of the 20th annual Boston University Conference on Language Development, Vol. 2* (pp. 816-827). Somerville, MA: Cascadilla Press.

- Veloudis, G., and Philippaki-Warbourton, I. (1984). *The subjunctive mood in Modern Greek*, Thessaloniki.
- Vendler, Z. (1957). Verbs and Times. *The Philosophical Review*, 66(2), 143-160.
- Voeikova, M., and Gagarina, N. (2002). MLU, first lexicon and the early stages in the acquisition of case forms by two Russian children. In M. Voeikova and W. U. Dressier (Eds.), *Pre- and protomorphology: Early phases of morphological development in nouns and verbs* (pp. 115-131). Munchen: Lincom.
- Weist, R. M., Lyytinen, P., Wysocka, J., and Atanassova, M. (1997). The interaction of language and thought in children's language acquisition: A crosslinguistic study. *Journal of Child Language*, Child-Language.
- Wexler, K. (1992). Some Issues in the Growth of Control In S. I. R. K. Larson, U. Lahiri, and J. Higginbotham (Ed.), *Control and Grammar*. Dordrecht: Kluwer.
- Wexler, K. (1994). Optional Infinitives, head movement and the economy of derivation in child grammar. In D. L. N. Hornstein (Ed.), *Verb Movement*: Cambridge University Press.
- Wexler, K. (1996). Development of the morphosyntax of verbal inflections. In M. L. Rice (Ed.), *Towards a Genetics of Language*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 106, 23-79.
- Wijnen, F. (1997). *Temporal reference and eventivity in root infinitives*. Paper presented at the Proceedings of the Workshop on the Interpretation of Root Infinitives and Bare Nouns - MIT Occasional Papers in Linguistics.
- Xanthos, A., and Gillis, S. (2010). Quantifying the development of inflectional diversity. *First Language*, 30(2), 175-198.

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A., et al. (2009). On the role of morphological richness in the early development of noun and verb inflection. *First language*.