Optional Structures in the Acquisition of Polish:

A Cross-Linguistic Perspective

Von der Philosophisch-Historischen Fakultät der Universität Stuttgart
zur Erlangung der Würde eines Doktors der
Philosophie (Dr. phil.) genehmigte Abhandlung

Vorgelegt von

Almut Klepper-Pang

aus Minden / Westf.

Hauptberichter:	Prof. Ian Roberts, PhD
Mitberichter:	Prof. Dr. Christiane Schaner-Wolles
	Prof. Dr. Grzegorz Dogil

Tag der mündlichen Prüfung: 2. Oktober 2003

Institut für Linguistik / Anglistik der Universität Stuttgart

2003

A Journey of Thousand Miles
begins with the First Step.
千里之行, 始于足下.

(Chinese Proverb)

To those who joined me on this journey
To my parents, my children Matthias and Julia and my husband Peter.

Contents

Chapter 1: **Introductory Remarks**...… 7

Chapter 2: **Reflections on the Course of Language Acquisition**………………. 13
2.0 Introduction……………………………………………………………. 13
2.1 Concepts of language: E-language vs. I-language………………………….. 13
2.2 The Innateness Hypothesis and the Poverty of the Stimulus Argument…… 16
2.3 The Principles and Parameters Approach………………………………….. 18
2.4 Conclusion………………………………………………………………….. 21

Chapter 3: **The Nature of Early Child Grammar: Theories & Hypotheses**….. 23
3.0 Introduction……………………………………………………………. 23
3.1 Small Children's Small Grammars: The *Small Clause Hypothesis*.……….. 24
3.2 Functional Categories in Child Grammar: The *Full Competence Hypothesis* 27
3.3 *Optionality* in Child Languages: Properties of *Optional Infinitives*.………….... 30
3.4 Accounts of *Optional Infinitives*.……………………………………………. 35
3.5 Open Problems: Implications for the Research Questions of this study…….. 39
3.6 Conclusion……………………………………………………………………. 42

Chapter 4: **The Target Language: An Outline of Polish**………………………... 43
4.0 Introduction……………………………………………………………. 43
4.1 Brief Typological Sketch of Polish………………………………………….... 44
4.2 Basic Characteristics of Morphology………………………………………… 45
4.3 The Verbal Paradigm: Finiteness and Verbal Conjugation Patterns……….. 47
4.4 The Pronominal System and the *pro drop*-Status of Polish…………………. 51
4.5 Conclusion……………………………………………………………………. 53

Contents

Chapter 5	**The Study: Polish Acquisition Data**	54
5.0	Introduction	54
5.1	Methodology……	54
5.1.1	Design of the Empirical Study…….	55
5.1.2	Longitudinal Data…….	56
5.1.3	Spontaneous Speech Criteria……...	57
5.1.4	Transliteration and Data Annotation…….	58
5.2	The Database…….	60
5.2.1	The *Dagmara*-Corpus……	61
5.2.2	The *Anna*-Corpus……...	62
5.2.3	The *Aleksandra*-Corpus…….	64
5.3	Data Classification, Encoding and Evaluation.....	66
5.3.1	Sorting and Classification of the Data……...	66
5.3.2	Variables in the *Data Evaluation Scheme.*……	67
5.3.3	A Two-Dimensional System of Data Analysis……..	74
5.4	From Counts to Analysis: Research Questions…….	75
5.5	Conclusion…….	78
Chapter 6:	**Results of the Data Analysis**…….	79
6.0	Introduction……...	79
6.1	*Optional Infinitives* in Polish? – The Distribution of Finiteness……...	79
6.1.1	Starting Point of the Analysis……	79
6.1.2	Distribution of Finiteness in all three Corpora……..	80
6.1.3	The Interpretation of Main Clause Infinitives: Context Analyses…….	86
6.1.4	Summary……	91
6.2	A Glance at the Earliest Stage of Acquisition……...	93
6.2.1	Dealing with Early Speech Data: *Aleksandra*-Corpus……..	94
6.2.2	The Nature of Earliest Verb Utterances (One- and Two-Word Stage)…….	94
6.2.3	The Development of Subject-Verb Agreement…….	97
6.2.4	Mechanisms in the Early Formation of Grammar: Cross-linguistic Evidence for *Surrogate Verb Forms.*……	99

Contents

Chapter 6: Results of the Data Analysis (cont.):

6.3 The Realization of Subjects in the *Aleksandra*-Corpus…….. 102
6.3.1 The Distribution of Null- vs. Overt Subjects…….. 103
6.3.2 Subject Type: NP vs. Pronominal Subject…….. 106
6.3.3 Comparison with Adult Speech (Control-Group)……... 111
6.3.4 Deriving Syntactic Dependencies in the *Aleksandra*-Corpus……. 114
6.3.5 The Realization of Subjects: A Cross-Linguistic Comparison……... 120
6.4 Conclusion…….. 125

Chapter 7: Polish Child Data in a Cross-Linguistic Context…… 129

Appendix:…….. 137

Data Evaluation Scheme…… 138

Results of Data Evaluation (*Aleksandra*-Corpus, files 10-37):

Table 1, files *Aleks* 10-14…….. 140
Table 2, files *Aleks* 15-19…….. 142
Table 3, files *Aleks* 20-24…….. 144
Table 4, files *Aleks* 25-29…….. 146
Table 5 files *Aleks* 30-34…….. 148
Table 6, files *Aleks* 35-37…….. 150

References…….. 152

Zusammenfassung…….. 162

Abstract……..…… 166

Chapter 1: Introductory Remarks

Much recent research on first language acquisition has focused on the phenomenon of *optionality* in early child grammar. The nature of optionality in child grammars is an issue that is still far from being settled, as the following recent statement by Nina Hyams (in her talk at BUCLD, see Hyams (2001:34)) illustrates:

> In one form or another the issue of optionality has been at the forefront of research into grammatical development for many years, especially as regards the child's use of functional elements. While the adult language may require the overt appearance of D(et) elements such as subjects and determiners and I(nfl) elements such as auxiliaries and verbal finiteness, children freely omit them in their speech. Despite the apparent pervasiveness of optional rules in early grammar, optionality is rather mysterious. Why is a rule that is obligatory in the adult grammar optional for the child? Optional rules raise problems both from a linguistic-theoretic perspective and from a learnability perspective.
> Linguistic theory has moved away from the optional rules of the Standard Theory (Chomsky 1965). Within current theory, optional processes are ruled out by licensing principles or by economy considerations (Chomsky 1992).
> From the perspective of a restrictive continuity hypothesis, we have to ask why the child's grammar would allow such rules?
> Moreover, optionality raises issues of learnability: if rules which are optional for the child are obligatory for the adult, the move from the child to adult grammar runs into a potential subset situation.

This is the puzzle of any study within the framework of optionality.

The idea for the topic of this thesis came up during the "*Optional Infinitive Stage"* of first language acquisition studies themselves, so to say, i.e. when research around the *Optional Infinitive Phenomenon* was becoming more and more popular[1]. Following Wexler (1994), a number of studies have addressed, and are still addressing, different aspects related to the phenomenon of *optionality* in different child languages.

[1] I would like to thank Sten Vikner and Ken Wexler for encouraging me to tackle this project in the first place. Special thanks to Sten for passing his enthusiasm for linguistics on to me (and many others who attended his lectures), and to Ken for showing so much interest in my work which he accompanied with many helpful comments and inspiring discussion, and for sending me all the current research papers from MIT.
I would also like to thank my supervisors Ian Roberts, Chris Schaner-Wolles and Grzegorz Dogil for their support and encouragement, as well as for many fruitful discussions and comments.

A major case of optionality can be found in verbal inflection: there seems to be a stage (around the age of 2;0) in which two verbal forms appear to exist in declarative main clauses: the adult-like finite form and an *optional* (root) infinitive that is ungrammatical in the target language. This stage, which is generally referred to as *Optional Infinitive (OI) Stage* (Wexler (1994)), has been observed in a number of languages, for instance Dutch, English, German and French. On the other hand, there are languages (such as Italian, Spanish and Tamil) which do not show the *Optional Infinitive* phenomenon.

The aim of this thesis is to contribute to the cross-linguistic discussion on OIs by providing empirical evidence from Polish, a rich-agreement language, which has not yet been analysed in this respect. Essentially, each form of the verbal paradigm in Polish is distinctly marked for person, number, and, in some tenses, also gender. In contrast to the Germanic languages, the infinitive represents a highly marked form in the Polish paradigm. Therefore, the study of the OI phenomenon in Polish seems to be of both empirical and theoretical interest: in addition to extending the previous empirical findings to a Slavic language, it might also shed light on the theoretical cross-linguistic OI-generalization proposed in the literature.

This thesis is based on a longitudinal study of Polish child language data which I collected in Gdansk / Poland over a period of three years.[2] The aim of this empirical study was to collect naturalistic, conversational data (matching the *spontaneous speech* criteria) from three children:

Dagmara: recordings between the age of 2;2 and 3;2.
Anna: recorded from 1;11 to 2;11.
Aleksandra: recordings between the age of 1;4 and 3;3.

Of the three corpora, the third set of data, the *Aleksandra*-Corpus, is of special interest since it covers data from the very beginning of child language production, thereby providing an insight into the earliest stages of language development.

[2] I am greatly indebted to my Polish collaborators for all their commitment in making our project work – a complex task that seemed impossible to realize at first. I would like to thank the families who accepted the tough job of reliably recording their children's speech at regular intervals and then assisting in the diffcult and time-consuming process of transliteration and data annotation: my special thanks go to Agnieszka Mackiewicz, Beata Plucińska, Jarek Pluciński and Małgorzata Paczkowska. I would also like to thank Krystyna and Andrzej Szczudło for helping me to select the children and their families in Gdańsk. Finally, I wish to thank Anna Tomczyk and Magdalena Kita for the effort they spent on the transliterations as well as for their native speaker judgments on the data.

In a nutshell, the three corpora are analysed with respect to the following research questions:

1. Existence of *Optional Infinitives* / main clause infinitives and possible constraints on their appearance; distribution of finite vs. non-finite forms
2. The onset of verbal inflection and its development in the earliest stages of acquisition (*Aleksandra*-Corpus)
3. The realization of subjects and the distribution of null- vs. overt subjects; further differentiation with respect to *type* of subject involved (lexical/nominal vs. pronominal).

This thesis is organised as follows:

Following these introductory remarks, **Chapter 2** lays the foundation for the theoretical framework of first language acquisition (henceforth "L1") studies. In the first section, I will present the Chomskyan perspective which is adopted in this study, and outline his conception of language acquisiton that distinguishes between I(nternal)-Language and E(xternal)-Language. In the next section, I will follow the argument for the existence of an innate language faculty that is linked to the Poverty of the Stimulus Argument. Apart from introducing the basic concepts underlying this approach, I will then outline the implications of the Principles-and-Parameters approach for language acquisition.

Chapter 3 reviews some of the theories and hypothesis on the nature of children's early grammatical system, notably the structure of early clauses and the presence of functional categories. Having outlined the Small Clause Hypothesis and the Full Competence Hypothesis, we will look at some evidence for structure-dependent relations in early child grammar. Finally, we will consider the *Optional Infinitive* phenomenon and present some of the accounts that have been proposed in the literature for this apparent deviation of the child system from the target grammar.

In **Chapter 4**, we will outline some of the morphosyntactic properties of Polish, the target language of the children analysed in our study. In this brief outline, I will confine myself to those grammatical properties that are relevant to the acquisitional phenomena presented in

chapter 6. After a brief typological sketch of Polish, I will discuss some crucial morphological features, in particular the concept of a word *stem* that is realized differently in Polish, compared to English. The next section introduces verbal conjugation patterns in order to show the salient features of the infinitive form in relation to the rest of the paradigm. The chapter ends in an overview of the pronominal system of Polish which will turn out to be relevant in chapter 6.3, when the realization of subjects in Polish child language will be discussed.

The next chapter (**Chapter 5)** deals with the empirical foundation of this study, i.e. a longitudinal study of spontaneous speech data from three Polish children which I collected in Gdansk / Poland with the help of three Polish families. The first section will introduce the methodology used in this study and outline the different steps in the process of collecting *spontaneous speech* data (recording, transliteration, sorting and classification of the data).

I will then give an overview of my database which consists of three sets of data, the *Anna-*, the *Dagmara-* and the *Aleksandra-*Corpus. The next section deals with the procedure of data classification, encoding and, by presenting the *data evaluation scheme* which I developed for my three corpora, data evaluation. The final section continues this theme by discussing the relevance of the variables introduced in the evaluation scheme to the research questions to be addressed in this study.

Chapter 6 covers the most important results of the data analysis (both quantitative and qualitative analyses). The three major research topics outlined above will be dealt with in three different sections, each including a discussion part as well as a cross-linguistic section. Crucially, the results of each research topic will also be evaluated in comparison to studies from other languages.

We will begin by presenting the distribution of finiteness, ultimately targeting at the question if there is an *Optional Infinitive* stage in Polish or not (6.1). In addition to this quantitative analysis, the occurring main clause infinitives undergo a qualitative analysis: their interpretation is studied in the form of a context analysis in order to assess their intended meaning.

The next section will deal with the earliest stages of acquisition, the one- and two-word stage (6.2). The unexpected finding of *Descriptive Imperatives* (instead of main clause infinitives) among the very first verbal utterances is discussed with respect to its implications for the clausal architecture of early child grammar. In a cross-linguistic discussion, the Polish findings are confronted with data from other languages.

We then move to the third research question that focuses on the realization of subjects. After establishing the distribution of null- vs. overt subjects, I will take a closer look at the *types* of subjects that are found in the *Aleksandra*-data, and their distribution.

Finally, the collected Polish data are related to data from other languages in a cross-linguistic comparison, before the results are summarized in the final section of chapter 6.

In the general discussion in **Chapter 7,** all results will ultimately be interpreted in the light of the phenomenon of *optionality* in early child grammar (cf. Q 4).

Chapter 2: Reflections on the Course of Language Acquisition

2.0 Introduction

In its long history, the study of language has frequently been understood as an inquiry into the nature of mind and thought on the assumption that "languages are the best mirror of the human mind", as the German philosopher Leibniz put it. According to the eighteenth century grammarian Beauzée, "the science of language does not differ at all from the science of thought".

These and other ideas of earlier traditions (see Chomsky (1966, 1986) for a review) re-emerged in the mid-1950s, with the development of what came to be called "generative grammar" ("where "generative" means nothing more than "explicit"", as Chomsky (1986:3) defines). Generative grammar is concerned with those aspects of form and meaning of language that are determined by the "language faculty", which is understood to be a particular component of the human mind. The general theory of linguistic structure that aims to discover the framework of principles and elements common to attainable human languages is now called "Universal Grammar" (UG). The focus of attention is knowledge of language: its nature, origin, and use, - spelled out in the famous three basic questions (Chomsky (1986:3)):

(i) What constitutes knowledge of language?
(ii) How is knowledge of language acquired?
(iii) How is knowledge of language put to use?

The Chomskyan perspective involved a shift of focus from behaviour or products of behaviour (language as a habit system) to states of the mind/brain that enter into behaviour.

The topic of this thesis is linked to the second question that has to be answered in terms of a specification of UG, i.e. a theory of the "initial state" of the language faculty, prior to any linguistic experience.

In the next sections of this chapter, we will review some of the concepts proposed in this framework. Section 2.2 deals with the distinction between Internal and External language.

2.1 Concepts of Language: E-Language vs. I-Language

According to Chomsky (1986:19-24), one can in principle study language either as an "external" or an "internal" phenomenon. Viewed "externally", the study of language involves no claim about the mental state of users of language "…in the sense that the construct is understood independently of the properties of the mind/brain." (Chomsky (1986:20). This approached is characterized as "E(xternal)-language".

Viewed "internally", however, the ability to speak and use a language is regarded as a mental capacity, and the study of language definitely involves claims about the mental states of speakers; that's why this view is characterized as "I(nternal)-language": "The I-language, then, is some element of the mind of the person who knows the language, acquired by the learner, and used by the speaker-hearer. … The grammar would then be a theory of the I-language, which is the object under investigation." (Chomsky (1986:22)).

Chomsky argues that I-language is both logically and epistemologically prior to E-language. E-languages are arbitrarily defined collections of linguistic objects (actions, utterances, linguistic forms); in general, there is no principled basis for choosing one collection over the other (e.g. corpus-based linguistics). In Saussurean structuralism, a language (*langue)* was taken to be a system of sounds and an associated system of concepts. To a large extent, our everyday conception of language also falls into the domain of E-language, as Roberts (1994) points out. Concepts such as "English" and "French", etc., are really E-language concepts referring to socio-political entities rather than mental entities.

On the other hand, all forms of E-language depend on I-language, in the sense that they are all the products of native speakers' linguistic capacities. I-language, then, can be characterized as the mental faculty that underlies the knowledge and the use of language. Some other mental faculties are involved in language use as well, of course, such as the perceptual system that facilitates comprehension, and the motor system responsible for language production. Besides, our beliefs and desires are usually involved, whenever we speak, which means that our system of beliefs (also called 'propositional attitudes') is activated, and yet this system of beliefs must be regarded as being *distinct* from our language. In this sense, Chomsky's earlier distinction between *competence* and *performance* (see Chomsky (1966)) can be reconstructed:

competence relates to I-language alone, while performance involves I-language in combination with other mental faculties.

I-language is taken to be the final state of a process of development of the language faculty from an initial state. The initial state of the language faculty is an innate endowment. According to Chomsky, the task of theoretical linguistics is the characterisation of the properties of I-language and the initial state. Theoretical linguists develop theories on different I-languages (that means, for instance, the I-language in the mind of a group of individuals referred to as “English speakers”). These theories are grammars (of “English”, “French”, etc.). The theory of the initial state is Universal Grammar (UG).

A theory of UG has to meet two requirements that seem to be incompatible at first sight: on the one hand, UG must be comprehensive enough to provide accounts for all kinds of grammatical properties of the world’s languages. UG should be able to explain how children develop *competence* in their native language so rapidly and apparently easily. Language acquisition suggests that, to a large extent, the final state is determined by the initial state. Thus, UG must have much in common with particular I-languages.

On the other hand, the observed diversity of the world’s languages poses a challenge for UG: it must be restrictive enough to allow for the attested variation, i.e. to allow the child to construct the grammar of his/her native language (which could, of course, be *any* of the approximately 6,000 existing world’s languages) from very limited and degenerate input (see section 2.2 for details).

Generative linguists take the aim of linguistics to be to account for the language faculty of human beings. To approach the nature of UG, linguists have to reconcile the two seemingly contradictory requirements mentioned above. The tension between these two requirements (being “comprehensive” and “restrictive” at the same time) can be resolved in the *Principles and Parameters* approach that will be discussed in section 2.3.
Ultimately, we are addressing the old question – often referred to as *Logical Problem of Language Acquisition* – of how come human beings possess linguistic knowledge of such richness, complexity and specificity, given the limitations of the data available. This argument will be continued in the following section when we discuss evidence for the view that the initial state of the language faculty – Universal Grammar – is an innate endowment.

2.2 The Innateness Hypothesis and the Poverty of the Stimulus Argument

This section deals with the well-known argument for the innateness of the language faculty (Innateness Hypothesis) (for a detailed discussion, see Atkinson (1992: ch. 3 & 10), Goodluck (1991: ch.1 & 4, Vikner (1995:7-10), Roberts (1994, 1997: ch. 5) and many others). The argument is based on the observed Poverty of the Stimulus to language acquisition.

Any approach to the nature of the language faculty involves an account of the shape of the linguistic knowledge (i.e. the grammar) in the brain (linked to Chomsky's question (i)) and also an explanation of how this knowledge enters the brain (corresponding to Chomsky's question (ii), see p. 9). The answer to the latter question determines possible answers to the former. As Vikner (1995:8) points out, there are three logical possibilities:

a. All properties of grammar are innate.
b. No properties of grammar are innate.
c. Some properties of grammar are innate, while some will have to be acquired.

The first possibility (a) can be refuted on the grounds of the following two simple arguments: if all properties of grammar were innate, we would expect all human beings to acquire and use the same language, which would equal the unrealistic concept of one common language for the human race.
Secondly, we would not be able to explain how languages can change over time.

The second possibility (b) would imply that the native speaker's mental grammar would have to be learned by means of general learning mechanisms, much like anything else in life that needs to be learned (mathematics for instance). It would then be surprising that young children below the age of six are so good at learning a language, while their general cognitive abilities as well as their learning skills are not very good at that age.

Another fact that would be unexplained under possibility (b) is the 100 percent success rate of language acquisition: human beings always learn their native languages perfectly (with the exception of pathological cases), and they do so extremely fast (compared to other learning processes like learning mathematics).

Yet every human language is extremely complex, as the following examples, taken from Chomsky (1986:8) demonstrate:

(1) I wonder who [the men expected to see them]

(2) [The men expected to see them.]

The clause bounded by brackets is included in both (1) and (2), but only in (2) may the pronoun *them* be referentially dependent on the antecedent *the men*. In (2) the pronoun is understood as referring in some manner indicated in the situational or discourse context, but not to *the men*. Numerous facts of this sort – subsumed under the theoretical notion of "binding theory" – are known without any instruction and without any relevant experience to differentiate the cases. How come - in cases like (1) and (2) - every child knows – unerringly! - to interpret the clause, i.e. the same string of words, differently in the two cases? How long has it taken linguists to study, say, the English languages with all its subtle complexities, and how far are we still from having a complete account? Yet, every single English child who is regularly exposed to English between the ages of one and four will acquire all the intricacies of the English language without any particular effort (and even without being explicitly encouraged or motivated to do so!)

This is all the more surprising when we consider the degeneracy of the direct linguistic input data to which the children are exposed. Three aspects are crucial here: the data are degenerate with respect to both *quantity* (the child never has a chance to hear all possible sentence and structures of the target language before her grammar is fully developed) and *quality* (not all of the sentences appearing in the child's input are well-formed, and the child has no clue to identify the ungrammatical input data). Thus, the input is deficient in two ways, which Roberts (1997:265ff) refers to as *incompleteness* and *noisiness of the input*.

The third aspect to be considered in this context refers to the fact that the child does not have access to negative evidence: children are usually not corrected every time they use a sentence that is not well-formed according to the rules of the target language. It was shown, in fact, that even in cases when they were corrected, these corrections did not make any difference, at least for the linguistic abilities of the children.

The idea that the stimulus to language acquisition is deficient in the three ways outlined above is generally referred to as the *Poverty of the Stimulus Argument*.

Another important fact that was revealed in comparative studies is the phenomenon of language universals. Despite the fact that there is an immense amount of variation between human languages, there is also a number of properties common to all of them (cf. Vikner (1995), Roberts (1997)). The existence of language universals would be a complete coincidence , if all properties of grammar had to be learned.

To summarize, we have seen evidence for possibility (c) above that part of the grammar of a native speaker is derived from innate knowledge and part is derived from the language spoken in the immediate environment of the child. This means that the language acquisition process is different from other learning processes. The claim made in (c) enables us to explain all the observations mentioned above: the 100 percent success rate despite the low age rate of the acquires, but also the speed and efficiency of the language acquisition process (especially in the face of the inaccessibility of negative input and the degeneracy of the data). The child acquiring his/her native language does not have to start from scratch, but already possesses some parts of the target/adult grammar. Language acquisition, then, results from the interaction of conditions within the learner (the initial state), and conditions outside the learner (the input). The less of the final state can be attributed to the input, the more it must be due to the initial state. The complexity observed in the final state (i.e. in the grammars of different languages), combined with the Poverty of the Stimulus Argument, leads us to conclude that a very significant part of our linguistic competence must already be present in the initial state, in other words: it must be innate.

2.3 The Principles and Parameters Approach

Having looked at the human language faculty that enables any child to acquire any native language (sometimes even two or three in parallel), we still have to account for the fact that there is a considerable amount of variation between human languages. This issue is addressed in the Principles and Parameters approach that is the subject matter of this section.

The Principles and Parameters approach is based on the idea that much of our linguistic competence – the principles of UG – are innate. What is open to variation are the parametric values that the principles can take on. Principles and Parameters theory evolved as a means of

resolving the tension between the two seemingly contradictory requirements imposed on UG: on the one hand, the need to posit a rich, comprehensive UG as a consequence of the Poverty of the Stimulus Argument, and on the other hand the variation which can be observed among the world's languages.

Roberts (1997:271ff.) outlines the Principles and Parameters theory as follows:
Apart from the claim of a rich, innate language faculty as a species characteristic (see discussion above), he adds two further statements that need to be stressed:

- There are no racial or cultural biases towards given languages or language types. And
- There is clear evidence that a sentence (or syntactic structure) which is well-formed in one natural language L may be ill-formed in some other natural language L'.

The Principles and Parameters conception of UG allows for cross-linguistic variation by associating with the principles of UG a small number of parameters of variation. The idea is that a given principle may be able to manifest itself in slightly different ways, along minimally differing parameters. Consequently, principles can be related to different parametric values, and a given association of principles with certain parametric values gives rise to a particular grammatical system: an I-language. Constructing a theory of Principles and Parameters involves claiming what is invariant and what is variant in UG and predicting the dimensions for language typologies. With respect to language acquisition, it involves predicting what features might develop in the course of language acquisition.

Roberts (1997:dto.) elaborates on the implication of the Principles and Parameters Approach for first language acquisition. I will present his argument here because it makes the whole approach more explicit and links a number of cross-linguistic syntactic findings to the domain of parameter setting. Under the assumption that parameters are binary, Roberts considers them as the truth values of contingent statements about grammars. In this sense, the principles of UG would be the necessary truths about grammars. (3) gives an example of how a parameter should be formulated:

(3) AgrS has enough inflectional content to recover the features of *pro* in Spec,AgrSP. (true/false). [true → a null-subject language]

In line with common practice in semantics, truth values can be stated as members of the set {1,0}. If we take several such parameters like (3) together and assume an ordering, we can give the set of parameter values of a given language as a binary number *n* (see Roberts (1997:273f.) for the concrete parameter values for English). Thus, the task of acquisition consists of assigning truth values to each parameter, and, ultimately, of finding the correct strings of 1s and 0s – the binary number – which characterizes the target grammar. This binary number could be thought of as the grammar's bar-code.

One of the questions that arises concerns the nature of the initial state and the way in which successive grammars are triggered before the adult grammar is attained. Although the acquisition process could theoretically be idealised as being instantaneous, child language studies have shown that children pass through a number of stages before attaining adult-like competence. These intermediate stages can be thought of as interim grammars, which approximate but do not quite match the target grammar.

According to Roberts, there are at least three ways in which one can conceptualise the process of language acquisition, and these are not necessarily mutually contradictory or exclusive.

1. Parameter settings may be initially open and filled in by acquisition.
2. Parameter settings may initially correspond to a random setting and the correct setting may eventually be obtained by exposure to the input data.
3. Parameters may be pre-set to a default (or unmarked value) with the derivations from the maximally unmarked system being reached through acquisition.

If we indicate an open setting (neither 1 nor 0) by '*' and default values for parameters as 0, we can schematise these respective conceptions as follows:

(4)	a.	[. . . * * * * * * . . .]	{open settings}
	b.	[. . . 1 0 1 0 1 0 . . .]	{random settings}
	c.	[. . . 0 0 0 0 0 0 . . .]	{default settings}

There is a lot of ongoing debate as to which of the three possible L1 scenarios is most suitable for an account of language acquisition. We will get back to some of the arguments and counter-arguments later in this thesis. In any case, two issues play an important role: the first one is the question whether maturation plays any role in acquisition. It has been suggested

that certain principles of UG may not be available at early stages of acquisition (Borer & Wexler (1987)). Rather, they develop as the child grows, as part of the general process of maturation. This view is a little problematic in the sense that it implicitly allows child grammars that would violate "adult UG" (although they wouldn't violate "general UG", of course). Such a hypothesis would require very strong empirical support.

The second issue concerns the theory of markedness. The question is: what are the default settings of parameters, and is there a corresponding default grammar (of which (4 c) might be a representation)? The idea that parameters have a pre-set value is an attractive one, and it has been proposed by many studies, beginning with the pioneering work of Hyams (1986). This approach, however, is problematic in two respects, as I will show in section 6.3.5

One proposal that has been made is known as the *Subset Condition* (see Berwick (1985)). It is based on the observation that the set of well-formed sentences in some languages is the subset of that of other languages. Given the lack of "negative evidence" for language acquisition (see section 2.2), children must avoid falling into "superset traps", or positing a grammar which generates a superset of the target language, since they will never have any evidence to "retreat" to the target grammar. Accordingly, it may be that subset grammars – or parameter-values giving rise to subset grammars – are unmarked.

In the next chapters, I will link these reflections on the three options in (4) to the observed facts of child language acquisition. Crucially, the relationship between the recorded speech data of children (which represent a kind of E-language in the sense of the discussion in section 2.1) and our questions about the acquisition of I-language is very indirect and must be mediated by linguistic theory. In any case, the way in which children acquire the morphological system of their native language and the implications which that may have for syntax, seems to be a key factor in the whole debate.

2.4 Conclusion

In this chapter, I have outlined the theoretical framework for much current research in the area of language acquisition. The first section (2.1) dealt with the Chomskyan perspective which is also adopted in this study, and outlined his conception of language acquisition. Chomsky

distinguishes between E(xternal)-language and I(nternal)-language. The concept of I-language is particularly relevant to our discussion, since I-language can be characterized as the mental faculty that underlies the knowledge and the use of language. Starting from the old question of the *Logical Problem of Language Acquisition*, we followed Chomsky's argument for the existence of an innate language faculty that is supported by the *Poverty of the Stimulus* argument.

Having established the claim that the initial state of the language faculty – Universal Grammar – is an innate endowment, I then outlined the implications of the *Principles and Parameter Approach* for language acquisition. Following Roberts (1997), I discussed three possible scenarios for the language acquisition process (depending on open, random or default settings of parameters).

What kind of consequences for empirical studies of first language acquisition does this approach have? One thing is obvious: the relationship between the recorded speech data of children (which represent a kind of E-language in the sense of the discussion in section 2.1) and our questions about the acquisition of I-language is very indirect and must be mediated by linguistic theory.

Chapter 3: The Nature of Early Grammar: Theories & Hypotheses

3.0 Introduction

The last two decades have seen a development in the theory of first language acquisition that Rizzi (1998) describes as follows:

> In recent years, an increasing number of theoretical linguists has started to pay serious attention to language development. Why is it so? Decades of pioneering work by theoretically oriented psycholinguists have conclusively shown that language development presents a highly structured mix of elements of continuity and discontinuity with respect to adult systems. This mix is extremely attractive for the theoretical linguist. Continuity, which clearly is the prevailing factor, makes sure that developmental evidence will bear on the object of inquiry that the linguist cares about, the study of systems constrained and made possible by the human language faculty.
> Discontinuity is what makes development interesting for us: we can reasonably hope that development will allow us to see properties that are not immediately accessible to observation in adult systems, thus allowing us to identify and explore neglected areas of the grammatical space defined by Universal Grammar.

It is not surprising that this growing interest in language development coincided with the consolidation of the Principles and Parameters model of UG in the mid eighties (see section 2.3). Parametric models introduced a theoretical framework well adapted for the comparison of systems basically cast in the same mould, but diverging on some structurally well-defined points. The same methodology used for comparative studies of adult languages was successfully extended to child language.

In this chapter, I would like to review an area of comparative acquisition studies which has been extensively investigated over the last decade: the phenomenon of *optional infinitives* in early systems. Before turning to the structural properties of *optional infinitives* in section 3.3 and possible accounts of this phenomenon in section 3.4, I will consider aspects of children's clausal architecture by presenting two conflicting hypotheses regarding the structure of early clauses: the *Small Clause Hypothesis* (section 3.1) and the *Full Competence Hypothesis* (section 3.2). This chapter ends with a discussion of open problems and their implications for the research question to be addressed in this study (section 3.5).

3.1 Small Children's Small Grammars: The *Small Clause Hypothesis*

In his influential paper on "Syntax at age two: cross-linguistic differences", Colin Phillips introduces the discussion as follows (Phillips (1995:325)):

> There is a lot missing from the speech of a typical two-year-old. This in itself is of no great interest to the linguist – after all, there are lots of other things that two-year-olds aren't great at. What makes the two-year-old more interesting to the linguist is that there are striking regularities in what gets missed out where.

These "striking regularities" and their possible interpretation are the main focus of this chapter. What do typical child utterances at this age look like?

At around two years children start to combine words. Although the first multiword utterances have a telegraphic character, they are not a mere simplification of the adult language, although it might look like it at first sight:

(1)	a.	Papa have it.	(Eve, 1;6[1])	
	b.	Cromer wear glasses.	(Eve, 2;0)	
	c.	Marie go.	(Sarah, 2;3)	
	d.	Mumma ride horsie.	(Sarah, 2;6)	[Radford (1990)]

The data (1a-d) represent a well-known observation in the speech of English-speaking children. From their first word combination up to about three years, English children often produce sentences in which either the third person singular inflection *–s* or the past tense marker *–ed* is missing, which results in the verb surfacing as a bare or uninflected form. Other typical sentences are shown in (2), where an auxiliary (either the perfective *have* or the progressive *be*) is missing, as indicated in square brackets, and only the participle form of the verb is overtly expressed:

(2)	a.	Eve gone [has].	(Eve, 1;6)	
	b.	Eve cracking nuts [is].	(Eve, 1;7)	
	c.	Mike gone [has].	(Sarah, 2;3)	[quoted from
	d.	Kitty hiding [is].	(Sarah, 2;10)	Guasti (2002:106)]

[1] In line with common practice in first language acquisition studies, the age of the children studied is indicated according to the formula: years;months.days (where the indication of the latter is optional).

In children's earliest multiword utterances, modals and the copula *be* are also frequently absent, as (3 a-b) illustrate:

(3)	a.	That my briefcase [is]	(Eve 1;9)
	b.	You nice [are]	(Sarah, 2;7)

A similar case of omission of functional material is displayed in (4 a-c) where the dummy auxiliary *do* is missing from negative sentences (4 a-b) and questions (4 c):

(4)	a.	Fraser not see him.	(Eve, 2;0)
	b.	He no bite ya.	(Sarah, 3;0)
	c.	Where ball go?	(Adam, 2;3)

In summary, the following functional elements are usually absent in (English) children's early clauses, with the result that children's speech strongly resembles telegraphic speech (see Brown (1973)):

- Grammatical morphemes (e.g. third person singular *–s*, past tense *–ed*)
- Auxiliaries (perfective *have*, progressive *be*)
- Dummy *do*
- Copula *be*

The functional elements listed above share a common property: they all express the feature content of the I-node. Since (English) children rarely use these elements during the early developmental stages, Radford (1990) has proposed that early clauses lack the corresponding inflectional category IP. Under this view, the child grammar's syntactic representation merely includes the lexical category VP, reflecting the concept of "Small Children's Small Grammars"[2].

As Guasti (2002:107) points out, the VP hosts the verb and all its arguments. Hence, an early clause would be a projection of the lexical properties of the verb and thus encodes the thematic relationships between the verb and its arguments (2002:107, referring to Haegeman (1994: ch.1)).

[2] In fact, "Small Children's Small Grammars" was the title of a talk Andrew Radford gave at the University of Stuttgart in summer 1993.

The hypothesis that children's clauses are VPs is also called the *Small Clause Hypothesis*, a label that is supposed to emphasize the similarity between early clauses and some adult structures that have also been analysed as lexical projections of the predicate (see Stowel (1983), but also Cardinaletti & Guasti (1995) for a critical discussion of the notion of small clause in relation to language acquisition).

An example of an adult small clause is the verbal complement following a verb of perception (5 a). This complement cannot include functional elements (see (5 b, c), and its verb cannot be inflected for tense and agreement. Therefore, such clauses have often been analysed as bare VPs (examples from Guasti (2002:108)):

(5) a. I saw Mary eat an apple.
b. *I saw Mary have eaten an apple.
c. *I saw Mary could eat an apple.

Thus, according to the *Small Clause Hypothesis*, the early grammar is a lexical-thematic system, in which lexical items project according to the X-bar structure and in agreement with the Projection Principles (which states that lexical information is syntactically represented). Although functional categories are part of Universal Grammar (UG), their availability is subject to maturation. That means that - according to Radford (1990) – functional categories are programmed to emerge and become operative around the age of three only.

Evidence for the *Small Clause Hypothesis* comes mainly from English. If we look at data from typologically different languages (in particular rich agreement languages), we can observe acquisitional facts that clearly contradict an analysis along the lines of the *Small Clause Hypothesis.*

In languages such as Catalan, Italian and Spanish, for instance, clitics are placed to the left of finite verbs, but to the right of infinitives, as illustrated in (6b) for Italian (Guasti (2002:149)):

(6) a. Gianni lo mangia.
Gianni it eat-3SG
'Gianni eats it.'

b. Maria ha promesso di mangiarlo.
Maria has promised of eat-INF-it
'Maria has promised to eat it.'

Torrens (1995) and Guasti (1993/4) have found that children acquiring Catalan, Spanish and Italian consistently place clitics in the correct position (as in (6)), without ever misplacing them. This finding already points to the direction of the argument in the next section: cross-linguistic investigations do not support the view that the initial grammar includes no functional categories.

Instead, most researchers agree that functional categories are present even in the earliest multiword utterances, though they may not always agree on the question of how much functional structure should be attributed to the child's underlying grammatical system.

In the next section, I will review some evidence from languages other than English against the *Small Clause Hypothesis*.

3.2 Functional Categories in Child Grammar: The *Full Competence Hypothesis*

The *Small Clause Hypothesis* has had the merit of bringing children's earliest productions into the focus of recent linguistic and psycholinguistic theories, but its claims have proven to be too strong.

Cross-linguistic studies have shown that children acquiring Danish, Dutch, French, German and Swedish also produce a considerable number of finite clauses, examples of which are given in (7) below (data from Guasti (2002:109)):

(7)	a.	Kann ikke see. can not see '(I) cannot see.'	(Danish, Anne, 2;0)
	b.	Hij doet 't niet. he makes it not 'He does not make it.'	(Dutch, Hein, 2;4)
	c.	Dort bébé. sleeps baby 'Baby sleeps.'	(French, Daniel, 1;11)
	d.	Da is(t) er. here is he. 'He is here.'	(German, Andreas, 2;1)

Defenders of the *Small Clause Hypothesis* might still argue that for these very young children, finite verbs are unanalysed chunks, which would mean that finite clauses are bare-VP *small clauses* in which the finite verb stays in V, as non-finite verbs do. This claim amounts to a clear prediction with respect to the distribution and structural position of these *small clauses*. This prediction can be tested by application of the same procedure of comparative analysis that is used for the study of word order phenomena in adult language.

If we analyse the distribution of verbs with respect to negation, for instance, we can see that the children treat finite and non-finite verbs differently. Pierce, for instance, convincingly shows that children acquiring French place the verb correctly with respect to negation, depending on if the verb is finite or non-finite: they correctly place a finite verb before the negation (*pas*) and a non-finite one after the negation, as the following samples show (Pierce (1992)):

(8) a. **Pas** manger la poupée. (Nathalie, 1;9)
not eat-INF the doll.
'The doll does not eat.'

b. Elle roule **pas**. (Grégoire, 1 ;11)
it roll-3SG not
'It does not roll.'

This correlation between finiteness and verb position with respect to negation is captured in the following table that documents Pierce's robust finding for French (Pierce (1992:65-66)):

(9)

	Finite Verb	**Infinitive**
Verb - pas	**216**	**2**
pas - Verb	**9**	**122**

(Data from Philippe, Nathalie and Daniel (age range 1;8-2;3))

Pierce's finding has been replicated for German and Dutch (Weissenborn (1990)).
This clear correlation between finiteness and verb placement with respect to negation provides evidence for the claim that children distinguish finite from non-finite verbs. In conclusion, we can state that the early grammar must contain - besides the lexical projection VP, as postulated by the S*mall Clause Hypothesis* - the functional projection IP, to

accommodate raising of finite verbs. Children express the morphosyntactic distinction between finite and infinitive verbs in terms of verb raising in the same way as adults do (see (8 b)).

A similar piece of evidence comes from Verb-Second (V2-) languages like German, where Poeppel & Wexler (1993) find a striking correlation between finiteness and verb position: the German child Andreas seems to reserve second position for finite verbs and final position for infinitives (at least in this one file), as the following table shows:

(10)

	V/2	**V/E**	**Total**
Finite Verb	**197**	**11**	**208**
(%)	94,7%	5,3%	82,9%
Non-finite Verb	**6**	**37**	**43**
(%)	14,0%	86,0%	17,1%

German: data from Andreas (1 file, age 2;1)

[Poeppel & Wexler 1993:7]

Although not all studies on German child language obtain such a clear distribution (see, for instance Schaner-Wolles (1995/6) for Austrian German), it is generally agreed that finite and non-finite verbs are distributed differently in children's clauses with respect to other clausal constituents. This discrepancy is unexpected under the *Small Clause Hypothesis,* which holds that all verbs should behave alike – that is, stay in the VP and thus follow negation (French), or occur in clause-final position (German).

The distributional facts, however, seem to point to the presence of functional categories in the child's grammatical system, and the adult-like use of finite clauses ultimately supports the hypothesis that children have *full competence.*

This hypothesis is usually referred to as *Full Competence Hypothesis.*

3.3 *Optionality* in Child Languages: Properties of *Optional Infinitives*

In the previous section (3.2), we have seen evidence for the presence of functional categories, and for the *Full Competence Hypothesis* that manifests itself in the adult-like, correct use of finiteness with respect to negation (French) or verb position (German). The evidence presented in table (9) and (10) in the previous section, for instance, clearly speaks in favour of a *Continuity View* of language development (cf. Rizzi's statement in the introduction (3.0)). A *Continuity View* amounts to positing that children's grammars include the same functional projections as adults' grammars, and that the early clausal architecture of the child's grammatical system corresponds to the adult one.

This view is challenged by the phenomenon of *optionality* in early child grammar. A major case of optionality can be found in verbal inflection: there is a stage (around the age of 2;0) in which two verbal forms seem to coexist in declarative main clauses: the adult-like finite form and an optional (root) infinitive[3]. In many – but not all – languages children at an early age produce infinitival constructions as main clauses, which is ungrammatical in the target languages. At the same time, finite forms alternate with such MCIs.

The presence of main clause infinitives, or – in Wexler's (1994) terminology – *optional infinitives,* challenges the continuity view of language development. Since these clauses are not acceptable in the target language, they represent a *discontinuity*. The question that arises immediately in this context is: how can we account for these deviant structures?

Various researchers have argued that *Optional Infinitive* clauses originate from grammatical deficits due to lack of maturation of relevant grammatical principles (Rizzi (1994), Wexler (1998)). It is assumed that, although the principles in question are part of UG, they are under the control of a biological program that makes them available at given points of development. Only when these principles mature, do *optional infinitive* clauses cease to be an option in the early grammar. A maturational approach is compatible with a continuity view of language development in the sense that the differences between the child and adult systems are

[3] I will henceforth use the neutral label "main clause infinitive" (MCI) as a non-committed term for the phenomenon which is called *optional infinitive (OI), root infinitive (RI)* or *null aux pattern* under different theoretical assumptions.

assumed to be constrained by UG (see Borer & Wexler (1987), and also chapter 2.3 of this thesis for discussion of the aspect of maturation).

The *Optional Infinitive* phenomenon can be best approached from its cross-linguistic dimension.

Main clause infinitives are attested for a variety of early languages, for instance for German, Dutch, English, Swedish, Danish, Norwegian and French, but not for Italian, Spanish, Catalan and Tamil.

Main clause infinitives are morphosyntactically infinitive verbs, and as such, they do not raise to I. In the previous section, we have seen that infinitive verbs, unlike finite verbs, follow the negation (*pas*) in French (and also German and Dutch), as illustrated in (11) below:

(11) a. **Pas** manger la poupée. (Nathalie, 1;9)
not eat-INF the doll.
'The doll does not eat.'

b. Elle roule **pas**. (Grégoire, 1 ;11)
it roll-3SG not
'It does not roll.'

A German example of the finite/non-finite alternation from the *Andreas*-file is given in (12):

(12) a. Ich **hab** ein dossen (= großen) Ball.
I have- 1.SG PRES a big ball
'I have a big ball.'

b. Thorsten Ball **haben**. [Poeppel & Wexler (1993:5f.)]
Thorsten ball have- INF
'Thorsten ball have' → 'Thorsten has a ball.' ?

(12 a,b) do not only show the alternation between finite and non-finite verbs in main clauses in early German, but also a clear preference for the use of finite verbs in Verb-Second- (V2-) and non-finite verbs in Verb-End-/Verb-final-positions. In that sense, the data (12 a,b) represent the distributional pattern displayed in table 10 in the previous section.

There are still other syntactic environments where main clause infinitives and finite verbs show an opposite distribution. The initial position in the clause is a case in point. Unlike first position in a finite clause, first position in an *Optional Infinitive* clause in early Germanic languages cannot be occupied by a non-subject constituent.

Besides, clitic and weak pronoun subjects are incompatible with *Optional Infinitives.* French learners, for instance, use subject clitic pronouns, such as *je* 'I' or *elle* 'she' only with finite verbs (according to Pierce (1992), at a frequency rate of 96%, i.e. 605 out of 632 subject clitics in total occur with finite verbs), as in (13):

(13) a. Il est pas là. (Nathalie, 2;2)
He is not there
'He is not there.'

b. Elle tombe. (Philippe, 2;2)
she falls
'She falls.'

In contrast to finite verbs, French children use nonclitic pronouns, such as *moi* 'me', with both finite and infinitive verbs, as Pierce shows:

(14) a. Bois peu moi. (Daniel, 1;8)
Drink-1.SG PRES little me
'Me drink little.'

b. Moi dessiner la mer. (Daniel, 1;10)
me draw-INF the sea
' Me draw the sea.'

It is generally assumed that subject clitics and weak pronouns are licensed by the agreement feature. From this assumption and the distribution of pronouns in early French (and Dutch, according to Haegeman (1995)), we can conclude that the agreement feature is positively specified in finite clauses, but not in main clause infinitives. Having already argued for the presence of a functional projection, IP, we can interpret the data in (14) as evidence for the claim that I includes the agreement specification needed to license subject clitics and weak pronouns, or (in the framework of distinct projections for individual features) that the early child grammar includes AgrP.

Another property of MCIs is revealed in a cross-linguistic comparison: main clause infinitives show a limited distribution. They are found in declarative sentences, but not in *wh*-questions (see Weissenborn (1990), Rizzi (1994)).

Besides, lexical verbs, but not auxiliaries can show up in main clause infinitives; that means that children do not produce sentences with an infinitive auxiliary (Poeppel & Wexler (1993)), as in (15):

(15) *Marie avoir mangé la pomme. [not attested]
Marie have-INF eaten the apple.
'Marie has eaten the apple.'

An interesting result from Austrian German child data should be considered here. With respect to another group of functional verbs, i.e. modal verbs and the copula *be*, Köhler (1998:80), comes to different results in her analysis of the *Nico*-data. She reports a few cases of the copula *sein* 'be' and the modal verb *können* 'can' in the infinitive:

(16) a. Emilie Pierre noch sein. (Nico 13 (2;6.12))
Emilie Pierre still be-INF
'Emilie is still with Pierre.'

b. Picki baun können (Nico 10 (2;05.01))
Picki (= Nico) build-INF can-INF
'Picki can build (it).'

Summing up the basic properties of *optional infinitives / main clause infinitives* in a variety of different languages, we can state the following results of our cross-linguistic survey:

- MCI clauses are not introduced by non-subject XPs in Verb-Second- (V2-) languages.
- MCIs are incompatible with clitic and weak pronoun subjects (notably French and Dutch).
- MCIs occur in declarative sentences, but not in *wh*-questions.
- MCIs seem to be incompatible with auxiliaries.

Due to the distributional analyses discussed above, many researchers agree on the claim that a child who produces an *optional infinitive* still knows that:

- Finite and non-finite verbs are distinct in terms of verb movement, with infinitives not raising to I,
  and that
- Finite clauses include functional projections

In the light of the continuity / discontinuity considerations outlined at the beginning of this section, the occurrence of *optional infinitives* raises (at least) the following questions:

- What factors license *optional infinitives* in early grammatical systems?
- What is the structure of OI-clauses?
- How does the OI-phenomenon disappear from the early child grammar?

In the next section, I will address the question of *why* there seems to be an *Optional Infinitive Stage* and what factors might determine it. Without going too much into the details of the individual analyses, I would like to give an overview of the two major "families" of theories that have been proposed for this phenomenon. A very good survey of the variety of existing theories and hypotheses is given in Guasti (2002), which I will follow here in the first place, and Phillips (1995).

When we follow the discussion on the OI-phenomenon over the past decade, we are provided with some insights not only into the nature of *optionality*, but also into the development and change of linguistic theory that is reflected in the (changing) accounts of OIs.

3.4 Accounts of *Optional Infinitives*

The discussion in the previous section has shown that the phenomenon of *optional infinitives* centres around an apparent contradiction: on the one hand, young children of approximately two years of age can be shown to have good knowledge of the morphosyntactic properties of finite vs. non-finite verbs, as shown by the distributional facts in table 9 and 10 above. On the other hand, however, they do not seem to know that infinitives cannot be used in main declarative sentences. To explain this apparent deviation from the target language, researchers have elaborated two main "families" of theories.

Generally speaking, one group of theories claims that the occurrence of *optional infinitives* is due to the option of leaving some functional feature underspecified (*Underspecification Accounts*). The other set of theories analyses *optional infinitive* clauses as reduced structures that result from the option of truncating structures at different levels of clausal architecture. In this section, we will briefly look at the essential ideas of the two families of approaches and discuss their (dis)advantages.

As far as *Underspecification Accounts* are concerned, we will consider the *Tense Omission Model* (Wexler (1999)) first.
As mentioned above, the inflectional node I is positively specified for a bundle of features in finite clauses. Among those features are agreement and tense features that express finiteness. When a feature is positively specified, it is generally morphologically expressed. *Underspecification Accounts* hold that a feature that is usually present in a finite clause (e.g. tense) fails to be specified or is missing in a given syntactic representation. In this case, the morpheme expressing that feature cannot surface, and the syntactic processes that depend on this feature do not occur. The choice of leaving a feature underspecified must be optional, since children produce both finite and non-finite clauses during this particular developmental stage (hence the name *optional infinitives,* that was already coined in Wexler (1994) in a slightly different theoretical framework).

There are two types of underspecification theories, differing mainly in which feature remains underspecified in the child grammar: *number* (Hoekstra & Hyams (1998)) or *tense* (Wexler (1994)). The latter has been further refined in Wexler (1999), who claims that either tense or

agreement or both can be left unspecified. For simplicity, I will look at only one of these approaches, the *Tense Omission Model*.

In a nutshell, it says that an OI clause arises when the child leaves the tense feature underspecified in a given clausal representation. According to this model, children are aware of the so-called *Tense Constraint* which says that a main clause must include a specification of tense. To explain why children can omit tense from the clausal representation, Wexler (1999) proposes an account couched in recent Minimalist theory. Adopting the VP-internal subject-hypothesis, Wexler proposes the *Checking Constraint* to explain why the subject moves from Spec VP first to Spec TP and then to Spec AgrP:
Both Agr and T have a D-feature, which must be eliminated by being checked against the D-feature of a DP-subject that raises to Spec TP and Spec AgrP.
According to this constraint, Agr and T are endowed with an uninterpretable D-feature that needs to be checked, and thereby eliminated, by the interpretable D-feature associated with the DP-subject. Wexler (1999) further assumes that children's grammar includes the *Uniqueness Constraint*, stating that a subject-DP can check the D-feature of either T or Agr, but not both. To avoid violation of this constraint, children can leave T underspecified, removing TP from the clausal representation.

To account for the fact that OI clauses are merely optional (children also produce finite clauses that include both TP and AgrP), Wexler proposes another constraint that is supposed to be present in both children's and adults' grammar: *Minimize Violations*: Given two representations, choose the one that violates as few grammatical constraints as possible (see Guasti (2002:135ff.) for a detailed account of how the Tense-, Checking, and Uniqueness Constraint and Minimize Violations interact).

Wexler's use of the term *constraints* differs from its use in the Principles and Parameters model. In Wexler's account constraints are soft and violable, much as in *Optimality Theory* (OT), (see Legendre et al. (2000) for an OT account of *optional infinitives* in French).

To sum up, in the *Tense Omission Model* the underspecification of a feature leads to the absence of the projection hosting this feature.

An alternative account to the *Optional Infinitive* phenomenon, the *Truncation Model*, is proposed by Rizzi (1994) who assumes the *Axiom of Clausal Representation:* CP is the root of all clauses (finite and non-finite). This axiom amounts to saying that all clauses, including declaratives, have a uniform representation: they are CPs, even in non-V2 languages, and even if the CP layer is not filled by lexical material.

While the axiom is always operative in adults' grammar, it only applies optionally in children's grammar. Therefore, in children's speech, some clauses are less than CPs, as functional projections can be truncated below CP. The Truncation mechanism works as follows. Truncation operates at the top level of the structural hierarchy and strips away every projection above the truncation site. Projections from the middle of the tree, however, cannot be removed.

Optional Infinitives, also called *Root Infinitives*, are structures truncated below TP, that means, they are VPs. Truncation at the VP level is possible for non-finite verbs, because they do not raise to I. The fact that RI clauses are structures truncated below TP explains their restricted distribution. If a given projection is cut off from the clausal representation, none of the morphosyntactic processes involving this projection can apply. Returning to the issue of clitics and weak subject pronouns, we can say now that these elements must be licensed by Agr, and – consequently – cannot occur in an RI clause, where AgrP is missing.

Again, the technical details of the truncation mechanism are not so relevant to the research questions of this thesis. We should summarize the key arguments of the two approaches and rather look at the implications of both Wexler's and Rizzi's theories for an account of the children's grammatical systems.

Back to the question of the clausal architecture of early child grammar, one should emphasize that both the U*nderspecification* and the *Truncation Account* of main clause infinitives are based on the idea that tense is somehow deficient in children's grammar. In the former case the deficiency is traced back to the *Uniqueness Constraint* (stating that a DP subject can check the D-feature of only one functional projection), and in the latter case it is linked to the hypothesis that different nodes can count as the root of the clause in children's grammar. Under both the *Truncation* and *Tense Omission Models* an *optional infinitive* clause is a reduced structure that lacks certain functional elements. In this respect, both theories retain

some of the ideas expressed in Radford's *Small Clause Hypothesis* (some early clauses may lack functional projections), but they crucially differ from this hypothesis in claiming that functional categories are part of children's grammar from their earliest multiword utterances.

Finally, let us consider two alternative proposals that follow a different line of argument.

One alternative proposal by Boser et al. (1992) and Whitman (1994) tries to account for the occurrence of main clause infinitives by postulating the *Null-Auxiliary Hypothesis*. According to this approach, MCIs in early child utterances are interpreted as [auxiliary/modal + infinitive]- structures, in which the auxiliary or modal verb has been dropped:

(17) $[_{CP}$ ein Bär$_i$ $[_C$ e$_{j/i}$ Infl] $[_{IP}$ t$_i$ $[_{VP}$ da putzen] $[_{I0}$ t$_j$ Infl]]]
a bear there clean$_{inf}$
'A bear (wants to / does /will) clean there.'

Arguing in favour of the *Strong Continuity Hypothesis*, Boser et al. assume that the same syntactic representations underlie children's utterances as adults' utterances, with the difference that heads and specifiers can remain empty in early child grammar. Therefore, Boser et al. analyse German matrix infinitives as full CPs with a non-overt inflectional element (*null aux*) in COMP (due to the assumption that movement of the inflected verb is obligatory in early child grammar). The null auxiliary is conceived of as an empty pronominal category carrying the *phi*-features inserted in I^0. It obeys the general licensing conditions for *pro* that are based on Rizzi's (1991) WH-Criterion and his (1986) *pro*-drop parameter:

(18) a) *pro* must be in a SPEC-head configuration with an overt X^0 that shares its grammatical feature specifications.

b) A pronominal (non-trace) empty X^0 must be in a SPEC-head configuration with an overt XP that shares its grammatical feature specifications.

[Whitman 1994:285]

Due to (18 b), the null I^0 in C is licensed only if the subject has moved to Spec CP, thus allowing for identification of the features of the null aux. According to Boser et al. and Whitman (dto.), this generalization matches their empirical data exactly: the null aux-pattern

does neither occur with fronted non-subjects, nor in non-subject wh-questions (see, however, Schaner-Wolles (1995/6) for arguments and data against the null-aux-hypothesis).

Finally, let's briefly look at a different approach (Phillips (1995, 1996)) that argues for the view that main clause infinitives (RIs) are *finite*. Phillips' argument is based on a cross-linguistic study of early child data from many different languages which he analyses with respect to "regularities of omission in two-year-old syntax and morphology" (Phillips (1995:325)). Leaving the details of his analysis aside, we can summarize his conclusions as follows:

a. Root Infinitive clauses are not due to a deficit in syntactic or morphological knowledge.
b. Root infinitive clauses are fully represented finite clauses in which merger of the verb with inflection has been delayed.
c. The cause of the delay in merging the verb with inflection is difficulty with the process of *accessing* morphological knowledge, which is not yet an overlearned, automatic process for the child.

Phillips' approach is attractive, as it involves neither mis-setting of a parameter (and the inherent learnability problem) nor the necessity to postulate maturation. I will get back to his proposal when discussing the results of my own data analysis in chapter 6.

3.5 Open Problems: Implications for the Research Question of this study

Having seen cross-linguistic evidence for the *optional infinitive* phenomenon from a variety of different languages, we will now turn to the implication of these findings for the research questions of this thesis that will be derived in this section.

(At least) one open question deserves to be raised first, however. It is the same question I formulated already towards the end of section 3.3:

How does the *Optional Infinitive* phenomenon disappear again from the early child grammar?

Some researchers have appealed to the concept of maturation, a biological mechanism that underlies the development of certain features of biological systems (see Rizzi (1994) and Borer & Wexler (1987)). Maturation is likely to control some aspects of language development, for instance the fact that infants start to babble around the age of 6-8 months. According to the maturational view, a genetic program also controls the development of syntax and determines the timing by which components of UG become available to the child. Under this view, main clause infinitives occur because principles of UG have not yet matured. The course of maturation is relatively independent of experience, which is supported by the fact that whether or not a component of UG is available, depends almost exclusively on the biological schedule. It is, for instance, not until age three that the *Uniqueness Constraint* disappears from children's grammar. Equally, it is not before that age that Rizzi's axiom about the CP being the root of the clause is genetically programmed to become fully operative. Once this has happened, *optional infinitive clauses* disappear from children's speech.

The general problem with accounts along the lines of maturation is that a maturational view always represents a less restrictive point of view than the idea that all child grammars are subject to the same constraints as adult grammars. Under a maturational view, children's grammatical systems are implicitly allowed to (temporarily) deviate from UG, or even violate the adult UG (not the "overall" UG of all human languages, however). For this reason, one needs very strong empirical support to justify a maturational account.

To conclude, the phenomenon of *optional infinitives* is still far from being settled. Despite the fact that there are a lot of different theories and hypotheses on the nature of main clause infinitives, there are still some open questions and issues that have not yet been covered in a unified account.

We definitely need more cross-linguistic data for comparison in order to check, if the apparent deviation from the adult target language is in fact a general and consistent phenomenon across early languages.

The purpose of this study, therefore, is to contribute to the cross-linguistic discussion on *optional infinitives* by providing empirical evidence from Polish, a rich-agreement language, which has not yet been analysed in this respect. From a morphosyntactic point of view, Polish seems to be a very good test case for the claims about the *Optional Infinitive Stage*. Essentially, each form of the verbal paradigm in Polish is distinctly marked for person / number. In contrast to the Germanic languages, the infinitive represents a highly marked form in the Polish paradigm (see chapter 4). Therefore, the study of the OI phenomenon in Polish seems to be of both empirical and theoretical interest.

On the one hand, the empirical part of this study (see chapter 5 for a methodological outline of the empirical design) aims at extending the previous empirical findings to data from a Slavic language, by providing new child data that I collected according to the *spontaneous speech* criteria, which makes them a reliable empirical foundation for both quantitative and qualitative analyses.

On the other hand, the analysis of my Polish data in a cross-linguistic framework might also shed light on the theoretical concept of *optionality* in children's early grammatical systems. Thus, I will analyse the data according to the following research questions:

Research questions:

Q1. Do the Polish child language data show any evidence for an *Optional Infinitive Stage*?

Concretely:

a. What does the distribution of finiteness look like in the three corpora?

b. As for the non-finite verb forms in main clauses (= "main-clause infinitives (MCI)"), - in what contexts do these MCIs occur and what is their intended meaning?

Q2. What do the *earliest* stages of acquisition reveal about the clausal architecture of early child grammar?

More specifically:

a. What do the earliest verb utterances at the one- and two-word stage (*Aleks 1-9)* look like?

b. What kind of acquistional *mechanisms and strategies* can be deduced from the earliest part of the *Aleksandra*-data?

Q3. How does the realization of subjects develop in the *Aleksandra*-data?

Concretely:

a. What does the distribution of null- vs. overt subjects look like?

b. What *types* of subjects can be identified and how are these subjects types (notably NP- vs. pronominal subjects) distributed?

Q4. How could the results of Q1-Q3 be interpreted with respect to the phenomenon of *Optionality* in early child grammar?

The results of this analysis will be presented in chapter 6.

3.6 Conclusion

In this chapter, I have presented a variety of studies, theories and hypotheses in an area of comparative acquisition studies which has been at the forefront of research: the phenomenon of *optional infinitives* in early child grammar. In the first two sections of this chapter (3.1 & 3.2), I discussed two conflicting hypotheses regarding the structure of early clauses, the *Small Clause Hypothesis* and the *Full Competence Hypothesis.* By reviewing empirical evidence in the form of child language data from different languages, I interpreted and evaluated the findings with respect to the two hypotheses.

Section 3.3 dealt with structural properties, and section 3.4 with possible accounts of *optional infinitives.* In section 3.5, we have looked at some open problems and their implications for the research question to be addressed in this study.

Before turning to the database and the results of the analyses, I will give a sketch of the essential morphosyntactic properties of Polish, the target language of the children analysed in the study.

Chapter 4: The Target Language: An Outline of Polish

4.0 Introduction

In this chapter, I will give a sketch of the basic morphosyntactic properties of Polish, the target language of the children that were analysed in this study. I would like to emphasize, however, that this chapter is not supposed to provide a comprehensive or even complete account of Polish grammar, - this (probably impossible task) would, in fact, go far beyond the scope of this study. In this brief outline, I will instead confine myself to those grammatical and morphological properties of Polish that are relevant to the discussion of the acquisitional phenomena presented in chapter 6. To save space, I will only select areas and phenomena of Polish grammar that are considered to be useful or even necessary background information for the evaluation of the Polish acquisition data as well as subsequent analyses.

The first section (4.1) deals with a brief typological sketch of Polish. This survey of the major typological features of Polish is followed by a an outline of some basic characteristics of Polish morphology (4.2). In particular, I will compare the concept of a word *stem* that is realized differently in Polish, to its realization in English. Since my analysis of the data focuses on the acquisition of verb morphology, I will then present some verbal conjugation patterns in section 4.3 to show what the adult / target system looks like. Crucially, I will show that the infinitive that plays a major role in the data analysis, does indeed show salient features in relation to the rest of the paradigm.

The chapter ends with an overview of the pronominal system of Polish (section 4.4) which will turn out to be particularly relevant in chapter 6.3, when the realization of subjects in Polish child language will be discussed.

4.1 Brief Typological Sketch of Polish

Polish is an Indo-European language belonging to the Slavic branch. Within the Slavic branch, Polish belongs to the group of North Slavic languages which is further subdivided into East Slavic (Russian, Ukrainian etc.) and West Slavic languages (Polish, Czech, Slovak etc.) (for a more detailed survey see Comrie (1987:322ff.)).

Like other Slavic languages, Polish displays many typical characteristics of an inflecting or fusional type of languages, although it appears to be less close to the ideal model of an inflecting language than Czech or Slovak (cf. Lotko (1979) who mentions various domains in contemporary Polish where synthetic structures are replaced by analytic expressions). Polish displays a rich inflectional system, in which single grammatical morphemes combine several functions: case, gender and number in the nominal paradigm; person and number in verbal inflection. In general, the grammatical system of Polish closely parallels that of Russian.

Due to its rich inflectional system, Polish shows considerable freedom in word order. Apart from the standard predominant SVO order as represented in (1a), various deviations from this unmarked order are possible, as the examples in (1b-f), taken from Witkoś (1993a:291f.), show.

(1)	a.	Jan	czyta	książkę.	S V O
		Jan	read-3SG PRES	book-acc.	
		'Jan	reads	the / a book.'	
	b.	Jan	książkę	czyta.	S O V
	c.	Książkę	Jan	czyta.	O S V
	d.	Książkę	czyta	Jan.	O V S
	e.	Czyta	książkę	Jan.	V O S
	f.	Czyta	Jan	książkę.	V S O

Deviations from the basic SVO order generally serve the purpose of topicalization and - in combination with focal stress - special emphasis on pragmatically more salient elements of

the clause. For a more detailed analysis of word order phenomena see e.g. Willim (1986), Witkoś (1993 a & b) & Zabrocki (1990).

4.2 Basic Characteristics of Morphology

In the following sections on the Polish target grammar, I will only mention those features which are relevant to the present study. For a comprehensive description of Polish grammar the interested reader is referred to Brooks (1975), de Bray (1969) and Saloni & Świdziński (1987).
From a contrastive point of view, Fisiak, Lipińska-Grzegorek & Zabrocki (1978) gives a useful outline of basic differences between English and Polish morphology (see also Smoczyńska (1985:596f.) for a contrastive sketch).

A comparison of the morphological systems of English and Polish reveals various fundamental differences. Crucially, the two languages differ in the realization of the concept of a word *stem.*

In English, the stem constitutes the basic form of a lexical item, both in nominal and verbal paradigms. The form "stem + 0" usually performs numerous functions; e.g. the same stem can function as the basic form of noun and verb (e.g. *play, talk, drink*). As a result, any complex form of the type "stem + inflectional ending" occurs in opposition to the basic stem, e.g.
walk-s vs. walk, play-ed vs. play, dog-s vs. dog etc.). Generally speaking, the use of inflectional endings is much more limited than in the Slavic languages.

In Polish, however, (and in the Slavic languages in general) bare stems are rare, and the word stem can hardly ever occur alone. Even those grammatical forms which are functionally unmarked, e.g. nominative singular of nouns or the infinitive of verbs, have specific and in most cases unambiguous endings which clearly specify the grammatical category of the given word. Due to the rich inflectional system in Slavic languages, a single form of a paradigm

occurs in opposition to the whole inflectional pattern, rather than to the stem or the basic form only. Therefore, the concept of a word stem in Polish represents a mere *abstraction* which can be formed by cutting off the core element in various forms of a word from its possible endings.

This contrast has crucial implications for language acquisition studies. The difference in the morphological system of the target grammar in English and Polish forces us to look at morphological development in both languages from different points of view. While morphological acquisition in English requires the ability to *add* grammatical morphemes to basic forms / stems, Polish children have to acquire the ability to *replace* grammatical morphemes according to the rules of the target system. In Polish (like in many other Slavic languages) the learner is confronted with additional difficulty in mastering the morphological paradigms because there is a lot of stem allomorphy to be found both in verbal and nominal patterns.

In consequence, although Polish children (as Smoczyńska (1985:597) points out) can easily split a word form into a stem and an ending (as illustrated by their own morphological formations, differing from adult forms), the basic unit is a *word* rather than a *morpheme*.

This observation is additionally strengthened by fixed penultimate word stress (unlike Russian and Serbo-Croatian, for instance, where word stress is mobile). Furthermore, since most Polish endings are syllabic, their presence does not affect the prosodic shape of the word.

4.3 The Verbal Paradigm in Polish: Finiteness and Verbal Conjugation Patterns

Within the verbal paradigm of Polish, both finite and non-finite forms are marked for aspect. Finite forms are inflected for tense, mood person, number, and in some instances also gender.

Crucially, non-finite forms are unambiguously marked, as can be seen in table 1 below. This clear morphological differentiation between finite and non-finite verb forms is of central importance to the debate on the *Optional Infinitive* phenomenon, since many languages which are traditionally classified as OI languages do not show a clear and unambiguous morphological distinction with respect to finiteness. In German, for instance, the form of the infinitive (e.g. *gehen* ‘to go’) is identical to two finite verb forms indicating first and third person plural form in the present tense, *wir gehen* ‘we go’ and *sie gehen* ‘they go’.

The major aspectual distinction in Polish is between perfective and imperfective, but Polish also has an iterative aspect. Perfective verbs are generally taken to specify a completed situation, i.e. a situation which has a beginning, a continuation and a termination.

The imperfective aspect is neutral with respect to the completion of an action. The imperfective form is morphologically simpler for most verbs, and the perfective form is derived from it. In most cases, this is done by either prefixation (2) or suffixation (3):

(2) pisać / na-pisać 'to write' (impf./pf.)

(3) kop-a-ć / kop-nąć 'to kick' (impf./pf.).

For some aspectual pairs, both verb forms do not differ in morphological complexity, but only in the choice of the relevant suffix. Many verbs of this group show stem alternations in addition to the suffix alternation, e.g. otworz-y-ć / otwier-a-ć 'to open' (pf./impf.). Suffix alternation usually implies the shift of the verb to another conjugation type. For a limited number of verbs (about a dozen) the basic form is perfective and the imperfective form is derived from it. e.g. dać / dawać 'to give' (pf./impf.) or wstać / wstawać 'to get up' (pf./impf.).

There are three moods in Polish: indicative, conditional and imperative. The tense system is very simple, consisting of three tenses only: present, past and future. The tense distinction appears in the indicative mood only, and its full realization concerns imperfective verbs. Perfective verbs lack proper present tense forms, since the notion of perfectivity precludes the simultaneity of action with the moment of speaking. Therefore, perfective verbs have only past and future forms, but those forms which would formally correspond to present tense forms (as in the imperfective paradigm) are used to express futurity. In contrast to this, imperfective verbs need a separate analytic construction to express futurity (involving the appropriate form of the future tense auxiliary of 'to be' (będę 'I will' etc.) followed either by the past participle of the main verb or the infinitive). The choice of the future form (past participle which involves gender distinction vs. infinitive) is optional.

Let's now turn to verbal conjugation patterns.

Table 1 below gives the conjugation pattern of the present tense for the three main verb classes of Polish. It should be noted that each verb has at least two stems.
Stem I, called the "present tense stem", which is given in table 1, e.g. *pisz-*, occurs in present forms (both "true" present tense forms of imperfective verbs and "formal" present tense forms of perfective forms, which are in fact future tense forms.)
Stem II, called the "infinitive stem", appears in past tense indicative forms (including the past participle of the analytic future) and the conditional forms.

Stem selection poses a major difficulty in Polish conjugation. Many verbs require stem alternations *within* a given inflectional paradigm, e.g. jad- vs. jedz- in the first conjugation class as given in table 1. With respect to stem alternations, Tokarski (1973) distinguishes eleven groups of verbs, some of them being further subdivided. Even for "regular" verb groups it is not sufficient to know the infinitive and the present tense form in order to be able to predict all the alternations involved. This lack of transparency in the verbal paradigm, therefore, presents the learner of Polish with considerable difficulties.

Table 1 below illustrates the conjugation pattern of the present tense for the three main verb classes of Polish.

Conj.Cl.	**General pattern:**		**Examples:**		**[Table 1**: ***Present tense*****]**	
1.	**1. Sg.**	**-ę**	(ja)	pisz-ę	id-ę	j**a**d-ę
	2.	**-esz**	(ty)	pisz-esz	idzi-esz	jedzi-esz
	3.	**-e**	on, ona ono	pisz-e	idzi-e	jedzi-e
	1. Pl.	**-emy**	(my)	pisz-emy	idzi-emy	jedzi-emy
	2.	**-ecie**	(wy)	pisz-ecie	idzi-ecie	jedzi-ecie
	3.	**-ą**	oni, one	pisz-ą	id-ą	j**a**dą
	Infinitive:			**pisać**	**iść**	**jechać**
				('write)	('go')	('drive')
	Imperative (Sg.):			**pisz!**	**idź!**	**jedź!**
2.	**1. Sg.**	**-ę**	(ja)	mówi-ę	prosz-ę	ucz-ę
	2.	**-isz / -ysz**	(ty)	mów-isz	pros-isz	ucz-ysz
	3.	**-i /-y**	on, ona, ono	mów-i	pros-i	ucz-y
	1. Pl.	**-imy/-ymy**	(my)	mów-imy	pros-imy	ucz-ymy
	2.	**-icie/-ycie**	(wy)	mów-icie	pros-icie	ucz-ycie
	3.	**-ą**	oni, one	mówi-ą	prosz-ą	ucz-ą
	Infinitive:			**mówić**	**prosić**	**uczyć**
				('speak')	('ask')	('teach')
	Imperative (Sg.):			**mów!**	**proś!**	**ucz (się)!**

3.	**1. Sg.:**	**-m**	czyta-m	rozumie-m	powie-m
	2.	**-sz**	czyta-sz	rozumie-sz	powie-sz
	3.	**-0**	czyta	rozumie	powie
	1. Pl.:	**-my**	czyta-my	rozumie-my	powie-my
	2.	**-cie**	czyta-cie	rozumie-cie	powie-cie
	3.	**-j-ą /-dz-ą**	czyta-j-ą	rozumie-j-ą	powie-**dz**-ą
	Infinitive:		**czytać**	**rozumieć**	**powiedzieć**
			('read')	('understand')	('say')
	Imperative (Sg.):		**czytaj!**	**rozum!**	**powiedz!**

[Table 1 above was adapted from Miodunka & Wróbel (1986)]

It should be noted that the infinitive represents a highly marked form within the verbal inflectional paradigm of Polish. Its salience is both morphologically and phonetically determined. Therefore, the Polish infinitive is not very likely to constitute *the* basic form of the verbal inflectional paradigm, as has been claimed for many other languages, in particular with respect to the OI debate. Instead, the imperative singular, a very frequent form in early child utterances, is much more likely to function as a basic form within the verbal paradigm. In addition to its frequency in use one might also argue for this assumption in terms of morphological simplicity.

We will test these hypotheses against our Polish acquisition data which are presented in chapters 5 and 6.

In any case, the fact that the Polish infinitive constitutes a highly marked form and occurs in unambiguous opposition to all the other verb forms is extremely convenient in view of the acquisition data analysis, since it allows reliable conclusions as to the occurrence of *optional infinitives* in Polish.

4.4 The Pronominal System and the *pro drop*-Status of Polish

This section deals with the pronominal system in Polish, and – related to that – the *pro drop /* null subject-status of Polish. I will not go into the details of a syntactic account of pronouns in Slavic here, however. The table below is just supposed to serve as background information for the discussion in chapter 6.

To get an overview of the pronominal system, let's consider the following table (Table 2) that shows the whole paradigm of Polish personal pronouns in all cases, including full pronouns and clitic pronouns (in brackets):

Table 2: Personal Pronouns in Polish:
Strong Pronouns vs. Clitic Pronouns (in brackets)

		Nominative	Accusative	Genitive	Dative
Sg.	1.	ja	mnie (**mię**)	mnie (**mię**)	mnie (**mi**)
	2.	ty	ciebie (**cię**)	ciebie (**cię**)	tobie (**ci**)
	3.	on (m)	jego, niego (**go**)	jego, niego (**go**)	jemu, niemu (**mu**)
		ona (f)	ją, nią	jej, niej	jej, niej
		ono (n)	je, nie	jego, niego, go	jemu, niemu, mu
Pl.	1.	my	nas	nas	nam
	2.	wy	was	was	wam
	3.	virile:			
		oni	ich, nich	ich, nich	im, nim
		nonvirile:			
		one	je, nie	ich, nich	im, nim

(Source: Zagorska Brooks (1975:296ff))

Polish is traditionally classified as a pro drop-/ null subject-language, as the following quotation shows (taken from Feldstein & Franks (2002:64), see also Franks (1995) for a parametric approach to null subject phenomena in Slavic languages):

> Subjects of verbs are normally omitted entirely when they are not emphasized, which explains why the nominative case forms do not have enclitics (in a sense, an omitted, or zero pronoun, is the nominative case clitic form).
>
> (Feldstein & Franks (2002:64)).

As for the group of strong pronouns, their use in the adult system is illustrated in (4) – emphatic use - and (5) in sentence-initial position:

(4) On widział **ciebie**, nie **jego.**
´He saw YOU, not HIM.‘

(5) **Mnie** daje chleb, nie tobie.
Me-DAT gives (he) the bread, not You-DAT.
‘He gives the bread to ME, not to YOU.’

Strong pronouns also occur in prepositional phrases (6 a) (in the adult grammar), whereas clitic pronouns are ungrammatical in this environment, as shown in (6 b):

(6) a. Poszedł po **ciebie**
He went after you-acc.
‘He went (out) to look for you.’
b. ***Poszedł po **cię.**
He went after you-acc-**cl.**.

In general, clitic pronouns occur in unstressed positions in the nominative, dative and genitive. Polish clitics enjoy a great deal of freedom with respect to positioning, but there is one main restriction: they can never occur sentence-initially.
In clitic-clusters, we find a dative < accusative order.
While clitics are not the subject matter of this study, I would like to refer to Witkoś (1998) for a very comprehensive study on the syntax of clitics in the Minimalist framework.

4.5 Conclusion

The preceding brief outline of some relevant aspects of the morphosyntactic properties of Polish has shown that – from the point of view of first language acquisition – the situation for children acquiring Polish is far from simple.
In particular, the lack of transparency in the verbal paradigm, due to stem alternations in the first place, presents the acquirer of Polish with considerable difficulties.

We have also seen that the infinitive represents a highly marked form within the verbal inflectional paradigm of Polish. Its salience is both morphologically and phonetically determined. Therefore, the Polish infinitive is not very likely to constitute *the* basic form of the verbal inflectional paradigm, as has been claimed for many other languages, in particular with respect to the *Optional Infinitive* debate. I will analyse my Polish data with respect to the distribution of finiteness in chapter 6.1. In any case, the fact that the Polish infinitive constitutes a highly marked form and occurs in unambiguous opposition to all the other verb forms is extremely convenient in view of the acquisition data analysis, since it allows reliable conclusions as to the occurrence of *Optional Infinitives* in Polish.

Before turning to the actual results of the data analysis (chapter 6), I will present the empirical foundation of my study, as well as the process of data collection and transliteration, in chapter 5.

.

Chapter 5: The Study: Polish Acquisition Data

5.0 Introduction

In this chapter, I will present the empirical foundation of this thesis, i.e. a longitudinal study of spontaneous speech data from three Polish children which I collected in Gdansk / Poland. Section 5.1 introduces the methodology used in this study and outlines the different steps in the process of collecting *spontaneous speech* data (recording, transliteration, sorting and classification of the data). Section 5.2 gives an overview of my database which consists of three sets of data, the *Anna-*, the *Dagmara-* and the *Aleksandra*-Corpus. Section 5.3 deals with the procedure of data classification, encoding and, by presenting the *data evaluation scheme* which I developed for my three corpora, data evaluation. Section 5.4 continues this theme by discussing the relevance of the variables introduced in the evaluation scheme to the research questions to be addressed in this study.

5.1 Methodology

This section outlines the methodological design of this study that is based on an analysis of tape-recorded data from three monolingual Polish children between the ages of 1;4 and 3;3. Section 5.1.1 begins by sketching the motivation for the design of this empirical study. The first important decision to be made concerns the *type* of data to be analysed. The choice of *longitudinal* speech data (as opposed to cross-sectional data) for the present study is discussed in section 5.1.2. A crucial feature of this longitudinal database consists in the nature of *spontaneous speech* data which is presented in section 5.1.3. Apart from introducing the key features of *spontaneous speech* data, we will also discuss the implications of these features for the process of data classification in first language acquisition studies. Finally, section 5.1.4 deals with how the *spontaneous speech* criteria are implemented in the process of transliteration and data annotation in the present study.

5.1.1 Design of the empirical study

This dissertation draws on a longitudinal study of spontaneous speech which was carried out in Gdansk / Poland. The aim of this empirical study was to collect naturalistic, conversational data from three children:

Aleksandra: aged 1;04 at the outset of the study
Anna: aged 1;11 at the outset of the recordings
Dagmara: aged 2;02 at the outset of the study.

First of all, two methodological issues are worth mentioning here: the *number of subjects* studied and the choice of *longitudinal* data.

As for the former, I decided to attempt to collect data from *more than one* child in order to overcome the notorious drawback of many L1 studies that are based on data of a single child. In general, the diversity of child data sources decreases the likelihood of errors in the data due to mistakes during collection and transcription. It also supplies a statistically good sample of data by minimizing the potential error rate due to reliance on any one child who might show retarded linguistic development (due to pathological reasons, for instance).

Another general problem in the collection of child language data has to do with the fact that – at the outset of the recordings - naturalistic child language data are unpredictable with respect to their usability for later analyses. Apart from children showing retarded linguistic development, there is generally a great deal of variation between individual children. It is a common observation that children differ not only in character, but also in talkativeness and answering behaviour during the recordings. As for the subjects in the present study, this inter-individual variation can be seen, for instance, when we compare the total number of analysable data and – even more crucially – the frequency of utterances containing a verb in all three corpora (as given in tables 1-3 in section 5.2 for every single file of each corpus). The degree of variation to be observed in the three corpora cannot simply be attributed to the age factor or different stages of language development, but points to inter-individual variation as an unpredictable variable in L1-studies.

Bearing these potential risks in mind, I decided to start recordings with three children, so that the future database would contain sufficient backup data, in case – as the recording were

progressing - one of the corpora should turn out to be less extensive or of less use for the analysis. At the same time, this wider sample of data would give us the chance to compare individual courses of language development.

On the other hand, one has to restrict the study to a manageable size of the database that is used for quantitative analyses. As far as my own database is concerned, it was impossible (and beyond the resources available to a single researcher) to analyse all three corpora (containing a total of 11,500 utterances) equally thoroughly.
Since the *Aleksandra*-Corpus represents the most extensive and revealing data collection covering the longest period of language development, I will focus primarily on these data in the quantitative analyses (based on my *data evaluation scheme,* as outlined in section 5.3). The data of the *Dagmara*- and *Anna*-Corpus will be used for specific research questions (in particular the distribution of finiteness) as well as for comparison and verification of conclusions drawn from a thorough analysis of the *Aleksandra*-data.

5.1.2 Longitudinal data

The decision to create and use a *longitudinal* database was due to a number of advantages and possibilities offered by this type of data.
First of all, longitudinal data allow us to start data collection at a very early stage of language acquisition, whereas experiments cannot really be done before the age of 2½ to 3 years (for the simple reason that, usually, children below that age are understandably not very cooperative in experimental settings). Longitudinal data, however, show the real course of language development, including variation in developmental patterns or unexpected deviation, thereby providing a reliable overall picture of the acquisition process.

Second, longitudinal data document long-term processes over a period of time, whereas cross-sectional or experimental studies tend to show isolated phenomena at a given point in development. This aspect can be crucial when dealing with data that are ambiguous or unclear (which is frequently the case in child language data!), since only longitudinal studies allow the researcher to compare these critical data with material from preceding or subsequent stages. Thus, even questionable or debatable findings that appear to lack a systematic

explanation, can be interpreted from a wider perspective if our analysis is based on continuous data.

Third, longitudinal studies show developmental trends in different areas of language acquisition, thereby revealing an interrelation between linguistic phenomena from different areas of language development (the correlation between null subjects, lack of finiteness and wh-contexts, which was found in *optional infinitive* studies, is a case in point).
Thus, the researcher might even been given the chance to reveal interesting correlations between phenomena that were not supposed to be related before.

Studies based on longitudinal data must meet one important requirement, however, in order to profit from the advantages outlined above: the data have to fulfil the criteria of *spontaneous speech* as established in Clahsen (1986).

5.1.3 Spontaneous Speech Criteria

The central idea behind the notion of *spontaneous speech data* is that language acquisition studies should be based on naturalistic data that represent the actual stage of a child's language development. Given that only genuine, spontaneous child language utterances can reflect the child's actual knowledge of the target language, all child data involving non-spontaneous structures, such as imitations or repetitions, should be excluded from the analysis. It is only via spontaneous speech data that language acquisition research can draw valid conclusions regarding the nature of the underlying child grammar (i.e. – in Chomsky's (1986) approach - the I-language).

As for the present data base, we set up the recordings as follows (according to the *spontaneous speech* criteria):
The recordings were usually carried out by the mothers, although I was present in some recording sessions at the beginning of the study, as well as later on at regular intervals. The recordings took place at home, i.e. in the normal environment of the children, where they were most likely to feel at ease. In addition to that, we paid special attention to the fact that only one person - usually the mother - was present during the recording. Before we started the

actual recordings, the mothers were introduced to the methodology of first language acquisition studies and the relevance of spontaneous speech data. In particular, they got to know the technical side of the recording process. During the first (test) recordings, I was present to train them a little and to discuss all sorts of "technical" problems that had come up during the test sessions. Since I could not be on the spot in Gdansk all the time, I wrote them a manual summarizing the main points as to how to proceed in the recordings and how to handle potential problems or difficult situations. To ensure the *spontaneous* nature of the data to be collected, I also listed a few things that should be avoided (although they seemed to be very tempting), such as correcting child utterances that were grammatically wrong, asking the child to repeat these sentences in the correct way, telling the child what to say or hindering her from expressing herself spontaneously. This manual was meant to be used as guideline during my absence so that the Polish collaborators made sure we would elicit child utterances that represented truly spontaneous speech, i.e. useful data for later analyses.

Recording sessions were often designed as play sessions in order to stimulate the child's language production. As the children grew older, we included various role plays, picture book descriptions and discussions about actual events in the child's direct environment. The variety of situational contexts during the recordings helped to stimulate utterances with different kinds of temporal reference, and not only utterances centring on the here-and-now context typically found in the earliest data.
The purpose of these measures was to ensure optimal conditions for gathering conversational data which can really be claimed to be both spontaneous and naturalistic, in other words: reliable data for the analysis.

5.1.4 Transliteration and Data Annotation

The final steps in the process of data collection involve data editing, i.e. transliteration, classification and data annotation.

The data (i.e. both the child's and the adult's/caretaker's utterances) were transliterated right after the recording in order to achieve reliable documentation. At the beginning of each transliterated file we gave a brief summary of the contents and the situation of the individual

recording, including some comments on the activities of the participants involved in the play sessions during that recording. More important still, we added contextual information to the individual child utterances in order to be able to assess the intended meaning. The fact that the recordings had been done by someone from the children's immediate environment turned out to be particularly advantageous, since the mothers (who were also part of the transliteration team) were most likely to infer the intended meaning as well as the intended temporal reference from the context, even in some critical cases where the interpretation was difficult. For any critical child utterance, we made a note of both linguistic and non-linguistic aspects that seemed to be relevant to the interpretation. As for the linguistic aspects, we would list grammatical features of the child utterances and – in case of ambiguous or deviant structures – the corresponding correct sentence in the target language. Non-linguistic notes would typically include information on the situational context of the utterance as well as a brief description of the child's actions while speaking. This method proved to be very useful for the context analysis of non-finite verb forms in matrix clauses, which will be presented in the next chapter.

After the transliteration, the data needed to be sorted and classified. An important first step involved checking every single utterances with respect to the criteria for spontaneous speech data as established in Clahsen (1986). Given that only genuine, spontaneous child language utterances can reflect the child's actual stage of development, Clahsen determines some diagnostic criteria for spontaneous speech data used in first language acquisition studies.

Applied to the three corpora of the present study, a number of utterances which did not meet these criteria had to be excluded from the analysis. First of all, a trivial case of unusable data is the set of utterances that are either incomplete or (partly) unintelligible. Such data are most frequent in early developmental stages, since they are often due to phonological deficiencies.

Less obvious cases of data that should be excluded are utterances that might look inconspicuous at first sight, but turn out to be non-spontaneous utterances once we examine the discourse in which they appear. In fact, it is necessary to check the context of every single utterance in order to identify different kinds of non-spontaneous data, such as utterances that are mere repetitions of the preceding dialogue or imitations of the preceding adult utterance.

Another kind of structure that should be excluded from the analysis are stereotyped or formulaic expressions or set phrases, such as the following examples (both very frequent structures in all three corpora):

(1) nie ma
not have-3SG
'There is none/nothing.'

(2) masz
have-2SG
'Here you are.'

These formulaic structures are very frequent in the children's input and tend to show up in the child language data from the early recordings on. It would, however, be rather rash to attribute an utterances like (2) to the grammatical knowledge of the second person singular agreement morphology. As a matter of fact, these formulaic expressions are usually taken to be unanalysed strings in the early child grammar, that is structures that are not necessarily due to grammatical knowledge, and are, therefore, excluded from the analysis.
Naturally, the same applies to forms of greeting, songs and quotations from fairy tales, books etc. that are due to rote learning rather than understanding and spontaneous wording.

This data classification process is followed by a systematic encoding and evaluation process that will be explained in section 5.3. Before turning to the issue of data evaluation, however, I would like to give an overview of the database in the next section.

5.2 The database

This thesis is based on a longitudinal study of Polish child language data which were collected in Gdansk / Poland over a period of three years. The aim of this empirical study was to collect naturalistic, conversational data from three children:

Dagmara: recordings between the age of 2;2 and 3;2.
Anna: recorded from 1;11 to 2;11.
Aleksandra: recordings between the age of 1;4 and 3;3.

All three children grow up in monolingual families in Gdansk in Northern Poland. Aleksandra and Dagmara are the only children in their families, whereas Anna has an older brother.

5.2.1 The *Dagmara- Corpus*

We started both the *Dagmara*- and the *Anna*-Corpus at an age which is traditionally claimed to represent the typical *Optional Infinitive* Stage, i.e. around the age of 2;0.
The recordings took place at intervals of 2-4 weeks; each recording lasted 45 minutes. The recordings were done according to the methodological design as outlined in section 5.1.

Table 1 summarizes the files that were analysed with respect to the distribution of finiteness, i.e. only the relevant recordings between the age of 2;2 and 2;8. The recordings (as well as the transliteration) were continued until the age of 3;2.

Table 1: ***Dagmara*-Corpus**

File	Age	Length	Utterances (total)	Analysable data	Utterances with verb	Frequency of verb utt.
1	2;02.25	45 min.	331	278	205	73,8 %
2	2;03.14	45 min.	293	261	142	54,4 %
3	2;04.3	45 min.	288	251	182	72,5 %
4	2;04.20	45 min.	286	245	197	80,4 %
5	2;05.13	45 min.	319	298	247	82,9 %
6	2;06.1	45 min.	316	273	212	77,7 %
7	2;07.3	45 min.	287	250	209	83,6 %
8	2;07.21	45 min.	293	255	224	87,8 %
		Total:	2413	**2111**	**1618**	**76,7% av.**

At the outset of the study, Dagmara was in the transitional phase between two-word- and multiword-stage. During the recordings, she did not show any inhibition to talk and, in fact, turned out to be a very talkative child – hence the high frequency of utterances containing a verb (76,7% on average, see rightmost column in table 1) testifying to an answering pattern in complete sentences rather than in single constituents. For the benefit of this study, Dagmara

produced a lot of truly spontaneous utterances, even without any stimulation from her mother or any other participants in the recording session. As a result, the number of utterances that needed to be excluded from the analysis for not meeting the *spontaneous speech* criteria is relatively low (compare the total number of utterances with the number of analysable data).

The results of the data analysis will be given in the next chapter.

5.2.2 The *Anna- Corpus*

Compared to the *Dagmara*-Corpus, the *Anna*-data represent an even earlier stage of acquisition since we started the recordings at the age of 1;11, when Anna was still at the two-word stage. Therefore, this corpus complements the *Dagmara*-data very well, in particular for an investigation of the *optional infinitive*-phenomenon in Polish child language, as discussed in the following chapter.

The recordings took place at intervals of 2-4 weeks; each recording lasted 45 minutes. Again, the recordings were carried out according to the same methodological principles as the other two corpora. In some of the recordings, there was one further member of the family present (taking part in the play sessions), to render the situation more natural, so that the child was more likely to feel at ease.

Table 2 below gives an overview of the files that were analysed with respect to the distribution of finiteness, i.e. only the relevant recordings from the age of 1;11 until 2;7. The recordings (as well as the transliteration) were continued until the age of 2;11. (see *Data Overview* with a list of all recordings in the appendix).

Table 2: ***Anna*-Corpus**

File	Age	Length	Utterances (total)	Analysable data	Utterances with verb	Frequency of verb utt.
1	1;11.23	45 min.	210	171	82	47,9 %
2	2;0.13	45 min.	242	212	70	33,0 %
3	2;01.3	45 min.	252	218	76	34,9 %
4	2;01.25	45 min.	233	202	92	45,6 %
5	2;02.26	45 min.	325	289	98	33,9 %
6	2;03.20	45 min.	282	251	89	35,5 %
7	2;04.13	45 min.	340	319	165	51,7 %
8	2;05.4	45 min.	260	231	72	31,2 %
9	2;05.23	45 min.	292	239	77	32,2 %
10	2;06.15	45 min.	276	260	124	47,7 %
11	2;07.6	45 min.	277	245	111	45,3 %
		Total:	2989	**2637**	**1056**	**40,1% av.**

In contrast to Dagmara, Anna is a rather timid girl, which made her feel a little inhibited to talk at the beginning. The frequency of utterances containing a verb is well below the corresponding frequency rates of the *Dagmara*-Corpus, as the rightmost column of table 2 shows (40,1% on average). This low percentage shows her tendency to respond to questions (even open questions) with single constituent-answers (mainly NPs), trying to avoid complete sentences. This low frequency of verb utterances makes the *Anna*-data the least extensive corpus in terms of *quality* (rather than quantity). She also uses a lot of imitative structures and repetitions that had to be excluded from a quantitative analysis for the reasons outlined in the previous section. Therefore, this corpus also provides comparably less analysable data in total. On the whole, the *Anna*-corpus contains less revealing data with respect to both quantity and quality, which led me to use this set of data primarily for the analysis of the acquisition of finiteness.

These results concerning the distribution of finiteness will be discussed in the following chapter.

5.2.3 The *Aleksandra*-Corpus

In addition to these two corpora, the third set of data, the *Aleksandra*-Corpus[1], is of particular interest, since it covers very early child language data from regular recordings (beginning at the age of 1;04., when Aleksandra was still at the one word stage).

During the first stages (one- and two-word stage), the recordings took place at short intervals of 1-2 weeks, whereas later recordings were done at 2-3 week intervals. With the exception of the initial sessions in the early stages, i.e. up to the age of 2;0 / Aleks15, each recording lasted 45 minutes (see table 3 on the next page for an overview).

In the earliest (and most difficult) phase of the study, the mother would break up the recordings into smaller sessions, depending on how communicative the child was. During these early stages, the mother also recorded a lot of spontaneous situations, i.e. whenever she felt the child was willing to communicate. Thanks to her sensitive way of handling the recordings, we are provided with a valuable, revealing and reliable documentation of the emergence of the very first verbs in Aleksandra's language development.

Thus, the *Aleksandra* data provide an insight into a very early stage of development, thereby completing the overall picture of the acquisition of Polish verbal inflection.

Since the *Aleksandra*-Corpus represents the most extensive and revealing data collection covering the longest period of language development, I will focus primarily on these data in the quantitative analyses. The analysis of this set of data, which includes more than 6,000 analysable utterances in total and over 3,200 verb utterances, will be based on the *data evaluation scheme* that I developed for this corpus. The process of systematic data evaluation according to this scheme will be outlined in section 5.3.

A detailed account of the results and conclusions drawn from a thorough analysis of the *Aleksandra*-Corpus will be given in chapter 6.

---

[1] Special thanks to Małgorzata Paczkowska for her kind cooperation over three years, especially her diligence in the recordings, her patience in the transliteration process, and finally her endurance that provided us with a continuous, uninterrupted documentation of her daughter's language development over two years. I am also grateful for her extensive child data annotation and detailed contextual information for the early recordings.

Table 3: ***Aleksandra*-Corpus**

	File	Age	Length	Analysable Data	Utterances with verb / in %
One-Word-Stage	1	1;04.15	30 min.	[*Note: counting not useful in the very early files 1-9*]	
	2	1;05.5	2 x 15 min.		
	3	1;05.14	15 min.		
Transition to Two-Word-Stage	4	1;06.8	15 min.		
	5	1;06.10	15 min.		
	6	1;06.14	15 min.		
	7	1;07.4	20 min.		
Two-Word-Stage	8	1;07.15	20 min.		
	9	1;08.7	20 min.		
Transition Multiword-Stage	10	1;09.14	20 min.	94	29 (=30,8%)
	11	1;10.0	25 min.	89	28 (=31,5%)
	12	1;10.13	30 min.	122	43 (=35,2%)
Multi-Word-Stage	13	1;11.9	25 min.	75	39 (= 52%)
	14	2;0.1	20 min.	61	31 (= 50,8%)
	15	2;0.15	30 min.	137	89 (= 64,9%)
	16	2;0.29	45 min.	273	143 (= 52,4%)
	17	2;01.21	30 min.	123	52 (= 42,3%)
	18	2;02.3	45 min.	297	153 (= 51,6%)
	19	2;02.20	45 min.	275	125 (= 45,5%)
	20	2;03.4	45 min.	314	161 (= 51,3%)
	21	2;03.18	45 min.	253	133 (= 52,6%)
	22	2;04.4	45 min.	228	114 (= 50%)
	23	2;04.18	45 min.	247	127 (= 51,4%)
	24	2;05.1	45 min.	213	115 (= 54%)
	25	2;05.15	45 min.	175	81 (= 46,3%)
	26	2;06.0	45 min.	236	137 (= 58,1%)
	27	2;06.16	45 min.	317	164 (= 51,8%
	28	2;07.0	45 min.	237	157 (= 66,2%)
	29	2;07.14	45 min.	268	165 (= 61,5%)
	30	2;08.0	45 min.	243	135 (= 55,6%)
	31	2;08.15	45 min.	221	126 (= 57%)
	32	2;09.7	45 min.	277	149 (= 53,8%)
	33	2;10.1	45 min.	228	132 (= 57,9%)
	34	2;11.17	45 min.	233	123 (= 52,8%)
	35	3;00.19	45 min.	221	114 (= 51,6%)
	36	3;01.24.	45 min.	293	179 (= 61,1%)
	37	3;03.2	45 min.	301	196 (= 65,1%)
		Total (files 10-37):		**6051**	**3240 (= 53,5% av.)**

5.3 Data Classification, Encoding and Evaluation

This section is supposed to outline the procedure of data classification, encoding and evaluation. Section 5.3.1. begins by explaining how the data are sorted and classified for later analyses. The quantitative analysis is based on the *data evaluation scheme*, which I developed for my Polish data. The variables introduced in the scheme will be presented in section 5.3.2. (see a copy of the *data evaluation scheme* in the appendix). The next section (5.3.3.) introduces the idea of a two-dimensional system of data analysis that was developed and used in the present database. The relevance of this data approach using interrelated variables to the research questions of this study will be discussed in the following section 5.4.

5.3.1 Sorting and Classification of the Data

A first step in sorting and classifying the data involves discarding all (partly) unintelligible or incomplete utterances. Second, all utterances that do not meet the spontaneous speech criteria (as discussed in section 5.1.3.) were excluded from analysis, leaving us with a remainder of intelligible and codable utterances. Within this set of analysable data, we sorted out all utterances containing a verb, irrespective of the verb type (modal verb, auxiliary, copula or main verb) or tense. Since this study focuses on the acquisition of finiteness, only verb utterances were taken into account.

As for sentence types, we mainly analysed declarative sentences and some intonation questions (spoken with rising intonation, but no change in word order involved). Subject- and object questions, however, were not included in the analysis, particularly as some very frequent structures seem to be used in a rather stereotyped or formulaic way, e.g.

(3) Co to jest?
What this is-3SG
'What is this?'

(4) Kto to jest?
Who this is-3SG
'Who is this?'

Formulaic structures of this kind were excluded since they do not represent truly spontaneous speech data conforming to the criteria discussed above (for details and further examples see section 5.1.3.).

Since imperatives turned out to play an important role in the early recordings, we captured commands – usually realized as imperatives – as well, but counted them as a separate category. The only imperative we excluded from the counts was the imperative of the verb 'to look', as given in (5):

(5) Zobacz
look-IMP
'Look!'

The imperative in (5) is more likely to function as a set phrase rather than a true command. It is often used as an insertion within an utterances rather than an independent utterance; - that's why it should not be analysed on a par with other imperatives.

Another set of data to be excluded from the analysis are forms of greeting or saying goodbye and other repetitive structures, such as songs or extracts from stories and fairy tales that are learned by heart rather than produced spontaneously.

Having sorted out all data that are to be analysed, all these utterances are numbered in order of occurrence in the transcript and then entered into the data evaluation scheme that will be discussed below.

5.3.2 Variables in the Data Evaluation Scheme

This section deals with how the child language data of this study are assigned to linguistic categories. The data evaluation scheme[2] that was used and developed for this purpose offers a uniform format to cover a great variety of data. In order to keep the investigation manageable within the scope of a single researcher, the data analysis only focuses on those aspects of grammatical development that bear on the research topics to be dealt with in this study.

[2] This data evaluation scheme for my Polish child data was developed following a similar one used in the L1 project "Erstsprachlicher Grammatikerwerb des österreichischen Deutsch im Vergleich" at the University of Vienna , run by Chris Schaner-Wolles to whom I am very grateful for many helpful suggestions. I changed the scheme, adapting it to the grammatical peculiarities of Polish, and extended it by the use of interrelated variables (see 5.3.3.). The scheme was restricted to variables that are relevant to the research questions raised in this thesis.

In my data evaluation scheme (which is given in the appendix in full length) I set up the following variable types and variable values:

Variable 1: Finiteness

The first variable type centres around the notion of finiteness, differentiating in the first place between utterances that contain finite and non-finite verbs as opposed to imperatives. Four variable values are introduced here:

1:	*imperative*	(treated as a separate category)
2:	*finite verb*	(further analysed in variable 2-7: verb type, verb morphology)
3:	*infinitive*	(analysis continued in variable 8: interpretation RI)
4:	*participle only*	(pursued in variable 6: future tense)

Imperative are counted in order to get an impression of their frequency in the overall set of verb utterances, but apart from that, they are treated as a separate category.

Finite verb utterances will be further analysed with respect to verb type (variable 2-3) and verb morphology (variables 4-7).

Infinitives will be pursued in variable 8 where we will look at their function as well as their meaning and interpretation.

Bare Participle (labelled *participle only*) will be picked up again in variable 6 where their potential future meaning is investigated.

Variable 2-3: Verb type

The next variable type only applies to finite utterances, as mentioned above. This second variable type consists of two variable categories (variable number 2 & 3) and specifies the verb type of the utterance, distinguishing between functional and lexical verbs.

Variable 2 covers modal verbs and auxiliaries that may be followed by an infinitive or a participle, which is spelled out (in terms of variable values) as follows:

1: *Modal verb only*	(modal verbs without an infinitival complement)
2: *Modal / auxiliary + infinitive*	(modals or auxiliaries followed by infinitive)
3. *Auxiliary + participle*	(auxiliary (być) + participle)

Variable 3 includes lexical verbs, the copula, and utterances with two lexical verbs of the type iść spać ('to go to sleep'):

1: *Main verb*	
2: *Copula*	(być 'to be')
3: *2 Lexical Verbs*	(type iść spać ('to go to sleep')

Variables 4-7: Verb Morphology

The third step in the classification of finite utterances the area of verbal inflection. Variables 4-7 encode different tenses: present, past and future tense. Variable 4 captures present tense morphology in different persons:

Variable 4: Present Tense Morphology finite verb	1: 1st pers. sing.
	2: 2nd pers. sing.
	3: 3rd ´pers. sing.
	4: 1st pers. plural
	5: 2nd pers. plural
	6 : 3rd pers. plural
	7: agreement error

The last variable value (4.7) marks agreement errors, i.e. cases where the child failed to provide the correct agreement morphology between subject and verb.

The variable values that are related to Variable 5 (*past tense* verb morphology) follow the same structure as the ones in Variable 4, i.e. explicit person and number distinctions:

Variable 5: Past Tense Morphology finite verb

1: 1st pers. sing.
2: 2nd pers. sing.
3: 3rd ´pers. sing.
4: 1st pers. plural
5: 2nd pers. plural
6 : 3rd pers. plural
7: agreement error
8: gender agreement error

The only difference – compared to variable 4 – is a differentiation between two kinds of agreement errors here: the usual agreement error referring to person and number agreement with the subject of the utterance (covered by variable value 5.7.) and in addition to that a gender agreement error (variable value 5.8.) that refers to the feminine/masculine distinction that is only found in past tense forms.

Variables 6 and 7 encode (periphrastic) future tense utterances. As there are two ways of expressing the periphrastic future in Polish, both of which are formed with the future tense of the auxiliary być ('to be'), one followed by an infinitive, and the other (stylistically superior) variant using a past participle instead.[3] Variable 6 is responsible for the distinction between these two options (variable values 6.1. and 6.2.). It also includes the possibility of bare future auxiliaries and bare participles:

Variable 6: Future Tense

1: aux + infinitive [analytic future tense]
2 : aux + past participle [analytic future tense]
3 : future auxiliary (e.g. będę 'I will' based on być-INF)
4: past participle only (elliptical): future meaning

All instances of "aux" in Variable 6 involve forms like będę 'I will' which are based on być-INF ('to be'). In the two complete analytic future tense structures (variables 6.1. and 6.2.), the choice of the form (infinitive vs. past participle) is optional, although the past participle construction appears to be better from a stylistic point of view.

[3] We will ignore another possibility that is related to the phenomenon of aspect: perfective verbs can express futurity with the help of present tense forms. In this study, however, we won't focus on aspect, and the variable "future" will always refer to the two analytic future tense forms (see chapter 4 for discussion and examples).

The auxiliary involved in the formation of the analytic future tense is inflected for person and number; this gives rise to the next variable that deals with the morphology of the future tense auxiliary:

Variable 7: ***Future Tense Morphology of the Auxiliary***

1: 1st pers. sing.
2: 2nd pers. sing.
3: 3rd ´pers. sing.
4: 1st pers. plural
5: 2nd pers. plural
6 : 3rd pers. plural
7: agreement error

Let's now turn to infinitives, one of the essential aspects of this study.
Variable 8 picks up all verb utterances that were classified as non-finite according to variable 1.3. above. Leaving formal aspects aside now, variable 8 involves a crucial interpretative step towards the meaning of these so-called *Root-* (or *Optional*) *Infinitives*. Based on the contextual information given in the transcripts, we assigned one of the following variable values to each infinitive:

Variable 8: ***Interpretation Root Infinitive (RI)***

1: modal / volitional
2: elliptical / context-licensed
3: clearly **non**-modal / **non**-elliptical
4: interpretation unclear

The first variable value refers to main clause/root infinitives (infinitives in main clauses where the infinitive is the only verb form of the sentence) that appear to have a modal or volitional meaning, i.e. the infinitive is most like to express a possibility, a wish or an intention. One should emphasize here, however, that the interpretation of non-finite structures is never easy since these structures lack any finite material. In this case, the meaning of the infinitive can only be deduced from the situation or the context of the utterance or from the preceding dialogue. This is where the aspect of data annotation (as discussed in section 5.1.4.)

becomes crucial: only with the help of carefully annotated data can we decide on the most probable interpretation of a matrix/root infinitive.

The need for contextual information also plays an important role in the next variable (8.2.) that stands for context-licensed infinitives, for instance elliptical structures (where the auxiliary is dropped but implicitly understood through the context or the preceding dialogue). These elliptical infinitives are often associated with future meaning (when following a question, for instance, as in (6) below, taken from the *Anna*-corpus:

(6)	Mother:	Co teraz będziemy robić?	
		What now aux.-FUT do-INF?	'What will we do now?'
	Anna:	Zdejmować koszulkę.	
		Take off- INF shirt- DIM.-AKK	'Take off the (little) shirt.'
			[*Anna* 6, 17 (2;3)]

Note that such elliptical answers would also be possible in the target language (even in English, see the English translation).

A truly ungrammatical use of matrix infinitives is found in variable value 8.3. that cover cases of root infinitives that are clearly **non**-modal/-volitional or **non**-elliptical, that is not licensed by the discourse.

Potentially ungrammatical uses of matrix infinitives are covered in variable 8.4. where a few unclear cases are included, i.e. cases where the interpretation of the infinitive was either not possible or inconclusive.

The next three variables 9, 10 and 11 take a closer look at the lexical content of subjects, direct objects and indirect objects. Variable 9 deals with how (overt) subjects are realized, differentiating between nominal and pronominal subjects along these lines:

Variable 9: Lexical content (overt) subject

1: Noun / NP
2: personal pronoun
3: other pronouns (e.g. demonstrative)

This variable classifies overt subjects according to the type of subject involved. The first variable value includes nouns, proper names and NPs, whereas variable number 9.2. and 9.3.

cover pronominal subjects, either personal subject pronouns (9.2.) or other pronouns such as demonstratives (9.3.).

Variable 10 then looks at how direct objects – usually realized in accusative case – are realized. Again, the first option listed is Noun / NP (variable value 10.1). The remainder of variable values cover pronominal objects. According to the pronominal system in Polish (which is outlined in chapter 4), we distinguish the following types of object pronouns (value 2-5):

Variable 10: Lexical content direct object (acc.)

1: Noun / NP
2: strong pronoun
3: clitic pronoun
4: demonstrative pronoun
5: reflexive pronoun (się/siebie)
6: missing pronoun
7: wrong pronoun

The last two variable values cover ungrammatical uses of pronouns, involving either omission of pronouns in contexts where they are obligatory (e.g. omission of the reflexive pronoun się, variable 10.6), or the choice of a wrong or inappropriate pronoun (variable 10.7.).

The same variable pattern is used for indirect or prepositional objects:

Variable 10: Lexical content indirect object / prepositional object:

1: Noun / NP
2: strong pronoun
3: clitic pronoun
4: demonstrative pronoun
5: reflexive pronoun (sobie/sobą)
6: missing pronoun
7: wrong pronoun

Variable 11 – the last variable of this data evaluation scheme - considers cases of indirect objects that are either realized as NPs (marked with dative case) or Preposition + NP, similar to the English *dative shift construction*, as in (7):

(7) a. He gave the dog a bone.
b. He gave a bone to the dog.

The *Aleksandra*-corpus contains a lot of data where the dative-NP that is supposed to mark an indirect object is replaced by a PP. Both cases are considered here and classified with respect to the object type involved – nominal vs. pronominal.

The results and exact figures of the data evaluation process are given in detail in the appendix, including a complete list of the counts for each file of the *Aleksandra*-Corpus from file *Aleks*10 on. As for the pronouns, I provided both the frequency of the pronoun type and some concrete examples for the individual pronouns used in the given file. The idea of these examples is to show the development in the use of pronouns as well as the order of occurrence.

5.3.3 A Two-Dimensional System of Data Analysis

Having established the individual variables of the data evaluation scheme, we should look at an additional, rather crucial feature of the evaluation system – its two-dimensional point of view. Each utterance is classified according to a combination of two variables: on the one hand, the variable type encoding properties related to finiteness and verb morphology (as described above), and on the other hand, the one related to subject properties, i.e. the distinction between Null- and overt subjects.

To put that in concrete terms, every utterance is assigned a variable value concerning its finiteness properties *plus* a variable value depending on if the subject is null or overt, i.e. [-subj.] or [+subj.] (see *data evaluation scheme*). In practice, this means that each utterance, once classified in one direction (concerning the grammatical properties of the predicate), is further differentiated with respect to an overt realization of its subject.

As a result, this method enables us to see already during the process of data classification if a given feature of the predicate strongly correlates with either null subjects or overt subjects, or – this is the third possibility – if there is a random distribution. To give a concrete example, one can check if any of the verb types subsumed under variable 2 and 3 appears to correlate with overt subjects and null subjects respectively.

As Polish is traditionally classified as a null-subject language (see chapter 4 for a discussion), which means that the subject is usually omitted (unless it is emphasized), one would not expect a high frequency of overt subject utterances in Polish child language. As a matter of fact, however, the Polish child data point to the opposite direction: there is a decisive percentage of utterances with an overt subject. If this percentage exceeds the corresponding frequency of overt subjects in the adult system, it will be interesting to investigate what factors might cause the excessive use of [+subj.]-utterances Polish child grammar. From this point of view, a data approach using interrelated variables seems promising: it might allow us not only to assess the frequency of overt subjects (which – on its own – would be surprising enough), but also to establish a correlation between certain grammatical properties of the predicate and the use of null- or overt subjects. This issue will be picked up again in the next chapter.

5.4 From Counts to Analysis: Research Questions

Having looked at the procedure of data classification, encoding and evaluation, we will now turn to discussing the relevance of the variables introduced in the evaluation scheme to the research questions to be addressed in this study.
One note beforehand: as mentioned above, the *data evaluation scheme* that I designed and used here makes no claim to be exhaustive. It is made for the individual needs of the present study that focuses primarily on the acquisition of finiteness and the distribution of null- vs. overt subjects.

The essential research questions of this thesis that were formulated in chapter 3 are repeated here for the sake of convenience:

Research questions:

Q1. Do the Polish child language data show any evidence for an *Optional Infinitive Stage*?

Concretely:

a. What does the distribution of finiteness look like in the three corpora?

b. As for the non-finite verb forms in main clauses (= "main-clause infinitives (MCI)"), - in what contexts do these MCIs occur and what is their intended meaning?

Q2. What do the *earliest* stages of acquisition reveal about the clausal architecture of early child grammar?

More specifically:

a. What do the earliest verb utterances at the one- and two-word stage (*Aleks 1-9)* look like?

b. What kind of acquistional *mechanisms and strategies* can be deduced from the earliest part of the *Aleksandra*-data?

Q3. How does the the realization of subjects develop in the *Aleksandra*-data?

Concretely:

a. What does the distribution of null- vs. overt subjects look like?

b. What *types* of subjects can be identified and how are these subjects types (notably NP- vs. pronominal subjects) distributed?

Q4. How could the results of Q1-Q3 be interpreted with respect to the phenomenon of *Optionality* in early child grammar?

The motivation for the design of the *data evaluation scheme* is closely linked to these research topics. If we want to make claims about the course of child language development, we need a solid empirical foundation to test our hypotheses on. Therefore, the basic aim of the scheme is to capture the relevant features of verb morphology in figures and not just in an impressionistic way. For any phenomenon to be observed in child language data, we need to check if these findings are representative and significantly frequent or just a vague tendency that occurs at random. In order to do this, one has to base the analysis on reliable counts and a data evaluation system that lends itself to quantitative analyses. This is especially true for research questions Q1a (concerning the distribution of finiteness) and Q3 a-b (targeting the

distribution of subject types etc.) that should be addressed in the form of quantitative analyses.

That does not mean, however, that all aspects of child language acquisition can or should be analysed quantitatively. In the case of the first research questions, for instance, the quantitative analysis of the distribution of finite vs. non-finite verb utterances (Q1a) must be complemented by a context analysis of the infinitives (Q1b), i.e. a *qualitative* analysis.

The second question (Q 2a-b), is another case in point: especially very early stages of acquisition, such as the early files of the *Aleksandra*-corpus (*Aleks* 1-9), do not lend themselves to a *quantitative* analysis (mainly due to the scarceness of analysable data in recordings at a young age). These very revealing early stages of acquisition should rather undergo a *qualitative* analysis, considering also contextual and situational factors as well as the general cognitive development of the child. The early *Aleksandra*-data will be analysed in this respect in chapter 6.

The *Aleksandra*-data from file 10 (age 1;9) on will be analysed according to the *data evaluation scheme* (against the background of research questions Q1a and Q3a-b).
Q3a-b. focus on the distribution of null- vs. overt subjects, which is analysed according to the evaluation method mentioned in section 5.3.3. The idea of using interrelated variables actually came up during the process of data analysis when I found a surprising percentage of utterances with overt subjects in neutral contexts (i.e. no emphasis involved). Given the fact that in Polish, subjects are normally omitted when they are not emphasized, the frequent occurrence of overt subjects in Polish child language is an unexpected finding that calls for an explanation. One way of approaching this phenomenon is to find out which variables the use of null- or overt subjects might depend on.

Therefore, this data approach using interrelated variables seems promising: on the one hand, it will allow us to assess the frequency of (optional) overt subjects as well as the *type* of subject involved, which is a crucial finding in view of the null-subject status of Polish. On the other hand, it might also allow us to establish a correlation between certain grammatical properties of the predicate and the use of null- or overt subjects.

The results of the data evaluation will be presented in the next chapter. All results will ultimately be interpreted in the light of the phenomenon of *optionality* (cf. Q 4) in early child grammar in the general conclusion in chapter 7.

5.5 Conclusion

To summarize, in this chapter I presented the empirical foundation of this thesis, i.e. a longitudinal study of spontaneous speech data from three Polish children. Section 5.1 dealt with the methodological design of this study and introduced the different steps in the process of collecting *spontaneous speech* data. After discussing the theoretical relevance of longitudinal studies and spontaneous speech data, I then gave an overview of my database that consists of three sets of data, the *Anna-*, the *Dagmara-* and the *Aleksandra*-Corpus (section 5.2). Section 5.3 outlined the procedure of data classification and encoding. Since the process of data evaluation involves a crucial interpretative step towards later analyses, it was necessary to introduce and explain the *data evaluation scheme* which I developed for my three corpora. This scheme determines how the child language data of this study are assigned to linguistic categories. In order to expose this crucial interpretative step as much as possible, the individual variables of this scheme were presented in detail and illustrated by examples. In Section 5.4, we discussed the theoretical relevance of the variables introduced in the evaluation scheme to the research questions to be addressed in this study.

Being familiar with the empirical and theoretical implications of the individual variables used in the data evaluation process, we can now turn to the actual results of the data evaluation.

Chapter 6: Results of the Data Analysis

6.0 Introduction

In this chapter, I will present the results of the data analysis (both quantitative and qualitative analyses) by following the research questions (Q1-Q3) in order of appearance in the previous chapter. Section 6.1 begins by addressing the first research topic that concerns the distribution of finiteness, ultimately targeting at the question if there is an *Optional Infinitive* stage in Polish. Apart from this quantitative study, the occurring main clause infinitives undergo a qualitative analysis in section 6.1.3, where their interpretation is studied in the form of a context analysis. Section 6.2 complements the previous analyses by looking at the earliest stages of acquisition, the one- and two-word stage (cf. Q2). The unexpected finding of *Descriptive Imperatives* (instead of MCIs) among the very first verbal utterances is discussed with respect to its implications for the clausal architecture of early child grammar in section 6.2.3. In a cross-linguistic discussion, the Polish findings are confronted with data from other languages. The issue of early grammar's clausal architecture and the presence of functional categories is continued in section 6.3 that deals with the realization of subjects in the *Aleksandra*-Corpus (cf. Q3). After establishing the distribution of null- vs. overt subjects, we will take a closer look at the *types* of subjects that are found in the data and their distribution. Finally, our findings are related to data from other languages in a cross-linguistic comparison in section 6.3.5. The results are summarized and discussed in relation to the phenomenon of *optionality* in the final section 6.4.

6.1 *Optional Infinitives* in Polish? - The Distribution of Finiteness

6.1.1 Starting Point of the Analysis

The first research question is linked to the *Optional Infinitive* debate that concentrates on the cross-linguistic observation of optionality in early child grammar. The debate on the phenomenon of main clause infinitives (henceforth "MCI"), which is presented in chapter 3, has been nourished by data from different languages, many of which, however, do not show

any clear morphological distinction between the infinitive and (some) finite forms. The situation looks different in rich-agreement languages, such as Polish, where the infinitive is a distinct form within the verbal paradigm. This is what makes the Polish data a good test case for the *Optional Infinitive* claim: the fact that the Polish infinitive constitutes a highly marked form and occurs in unambiguous opposition to all the other verb forms is extremely convenient for the acquisition data analysis, since it allows reliable conclusions as to the occurrence of *optional infinitives* in Polish.

Let's therefore address the first part of research question Q1, i.e. Q1 a:

Q1. Do the Polish child language data show any evidence for an *Optional Infinitive Stage*?

Concretely:

a. What does the distribution of finiteness look like in the three corpora?

b. As for the non-finite verb forms in main clauses (= "main-clause infinitives (MCI)"), - in what contexts do these MCIs occur and what is their intended meaning?

All three corpora will be analysed with respect to the distribution of finite vs. non-finite verbs. The results of this quantitative analysis according Q1a will be given in section 6.1.2. The other half of this first research question (i.e. Q1 b) that involves a *qualitative* analysis of the context of the MCIs will be addressed in section 6.1.3.

6.1.2 Distribution of Finiteness in all three Corpora

First of all, all three corpora of the data base are analysed with respect to finiteness in order to see if the phenomenon of *optional infinitives* could also be shown to appear in the acquisition of Polish. Table 1 and 2 of this section give the distribution of finite vs. non-finite verbs in the *Dagmara*- and *Anna*-Corpus. Since the imperative will be claimed to have a particular status, the occurrence of imperatives is listed in a separate column. The results of the distribution of finite vs. non-finite verb utterances (leaving aside imperatives) in the *Dagmara*- and *Anna*-Corpus are illustrated in Figure 1 and 2 respectively.

Table 1: Distribution of Finiteness: *Dagmara*-Corpus

File	Age	verb-utt.	Imp.	/ %	Fin.V	/ %	Infin.	/ %
1	2;02.25	205	3	1.5 %	181	88.3 %	21	10.2 %
2	2;03.14	142	6	4.2 %	114	80.3 %	22	15.5 %
3	2;04.3	182	8	4.4 %	133	73.1 %	41	22.5 %
4	2;04.20	197	11	5.6 %	153	77.6 %	33	16.8 %
5	2;05.13	247	19	7.7 %	219	88.6 %	9	3.7 %
6	2;06.1	212	13	6.1 %	196	92.5 %	3	1.4 %
7	2;07.3	209	11	5.3 %	197	94.2 %	1	0.5 %
8	2;07.21	224	9	4.1 %	214	95.4 %	1	0.5 %

As mentioned in chapter 5, Table 1 and 2 cover only those files that were analyses with respect to the distribution of finiteness, i.e. only the relevant recordings between the age of 2;02 and 2;08 (*Dagmara*-Corpus) and 1;11 to 2;07 (*Anna*-Corpus).

Table 2: Distribution of Finiteness: *Anna*-Corpus

File	Age	verb-utt.	Imp.	/ %	Fin.V	/ %	Infin.	/ %
1	1;11.23	82	7	8.5 %	71	86.6 %	4	4.9 %
2	2;0.13	70	3	4.3 %	67	95.7 %	0	0 %
3	2;01.3	76	2	2.6 %	73	96.1 %	1	1.3 %
4	2;01.25	92	7	7.6 %	82	89.1 %	3	3.3 %
5	2;02.26	98	6	6.1 %	85	86.7 %	7	7.2 %
6	2;03.20	89	2	2.3 %	86	96.6 %	1	1.1 %
7	2;04.13	165	5	3.1 %	157	95.1 %	3	1.8 %
8	2;05.4	72	3	4.2 %	68	94.4 %	1	1.4 %
9	2;05.23	77	1	1.3 %	74	96.1 %	2	2.6 %
10	2;06.15	124	4	3.3 %	118	95.1 %	2	1.6 %
11	2;07.6	111	2	1.8 %	107	96.4 %	2	1.8 %

Leaving imperatives aside now, the distribution of finite vs. non-finite verb utterances in both corpora is illustrated in Figure 1a and 1b below. In Figure 1a, the distribution of finiteness is shown in absolute numbers (corresponding to the figures in Table 1) for the *Dagmara*-data:

Figure 1a: Distribution of infinitives vs. finite verb utterances (*Dagmara*-Corpus):

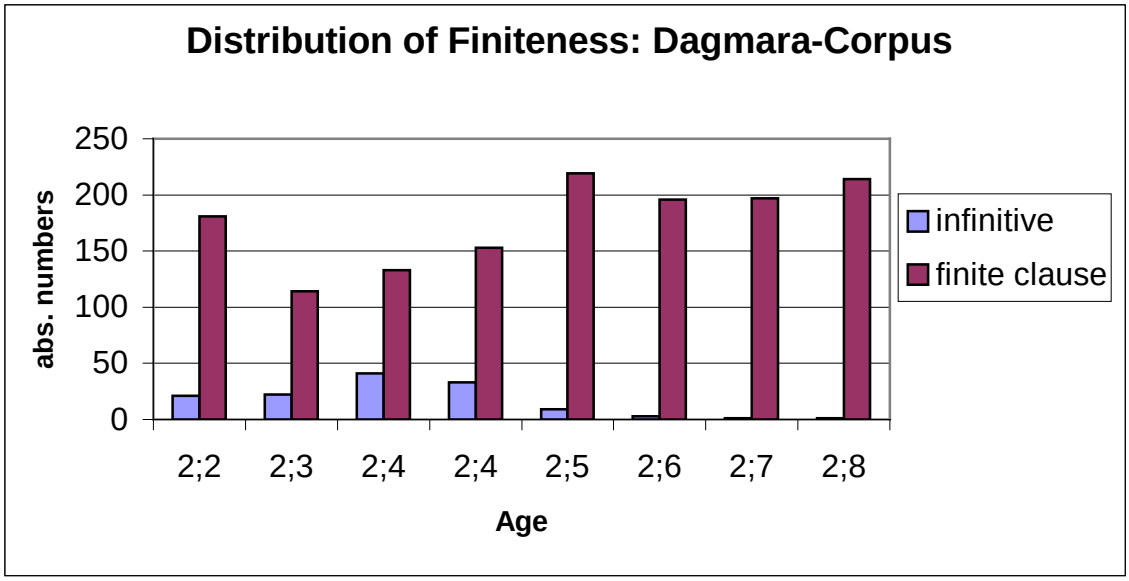
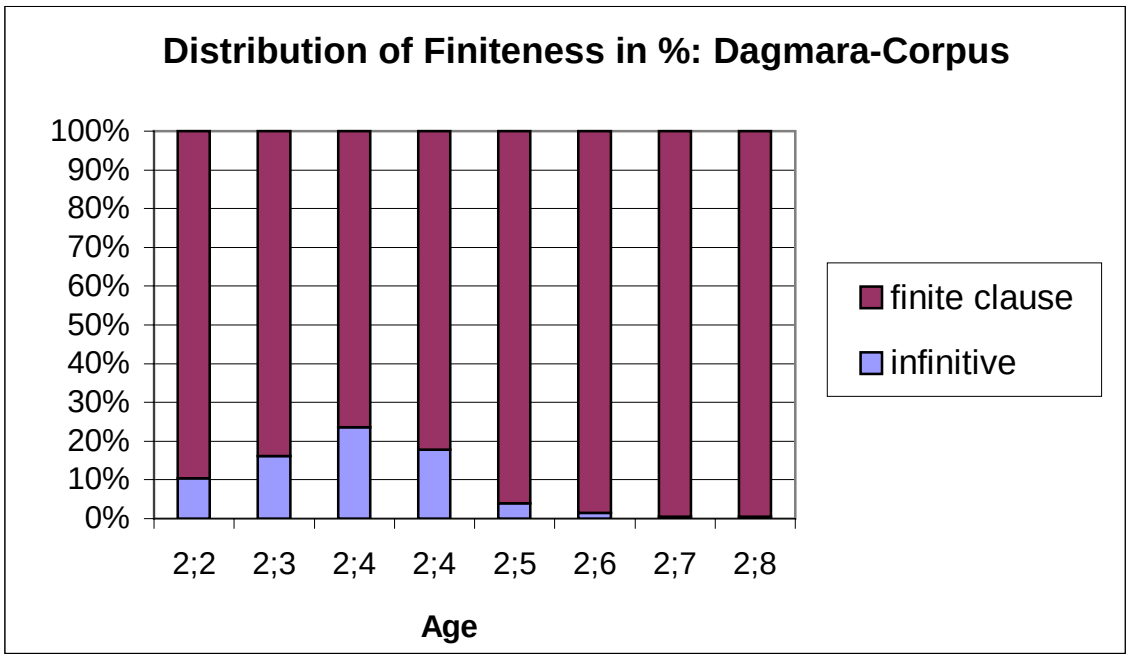


Figure 1b below illustrates the distribution of finiteness in the *Dagmara*-Corpus in percentage terms:

Figure 1b: Distribution of infinitives vs. finite verb utterances in % (*Dagmara*-Corpus):



Figures 2a and 2b illustrate the distribution for the *Anna*-Corpus. Based on the figures in Table 2, Figure 2a below covers the occurrence of infinitives vs. finite clauses in absolute numbers. Imperatives are excluded from this diagram:

Figure 2a: Distribution of infinitives vs. finite verb utterances (*Anna*-Corpus):

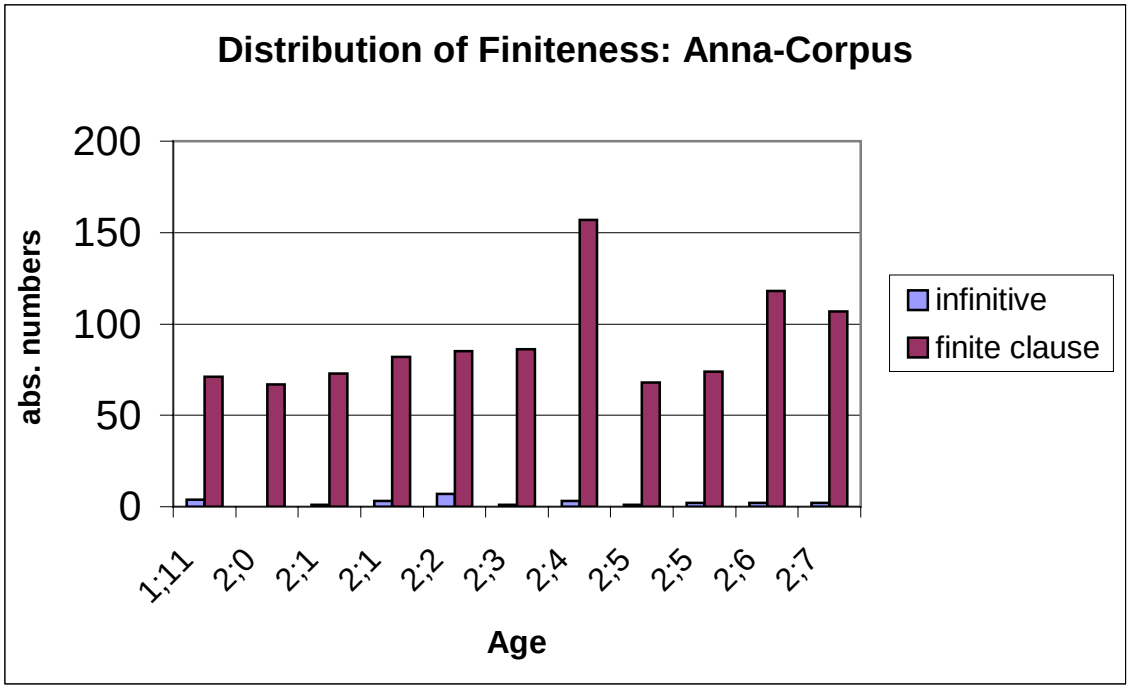
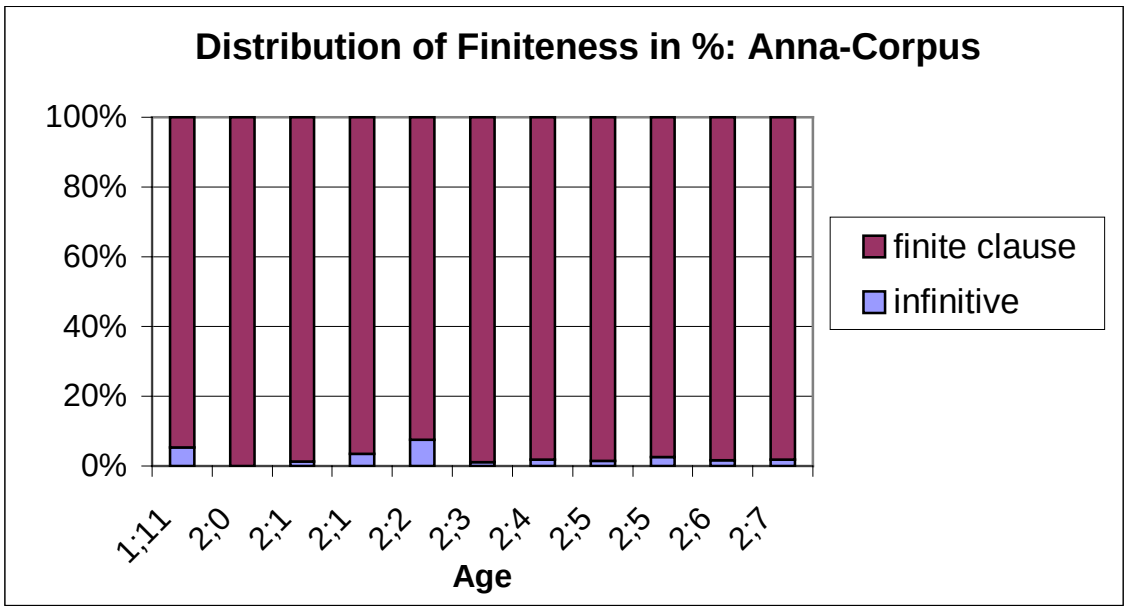


Figure 2b below illustrates the distribution of finiteness in the *Anna*-Corpus in percentage terms:

Figure 2b: Distribution of infinitives vs. finite verb utterances in % (*Anna*-Corpus):

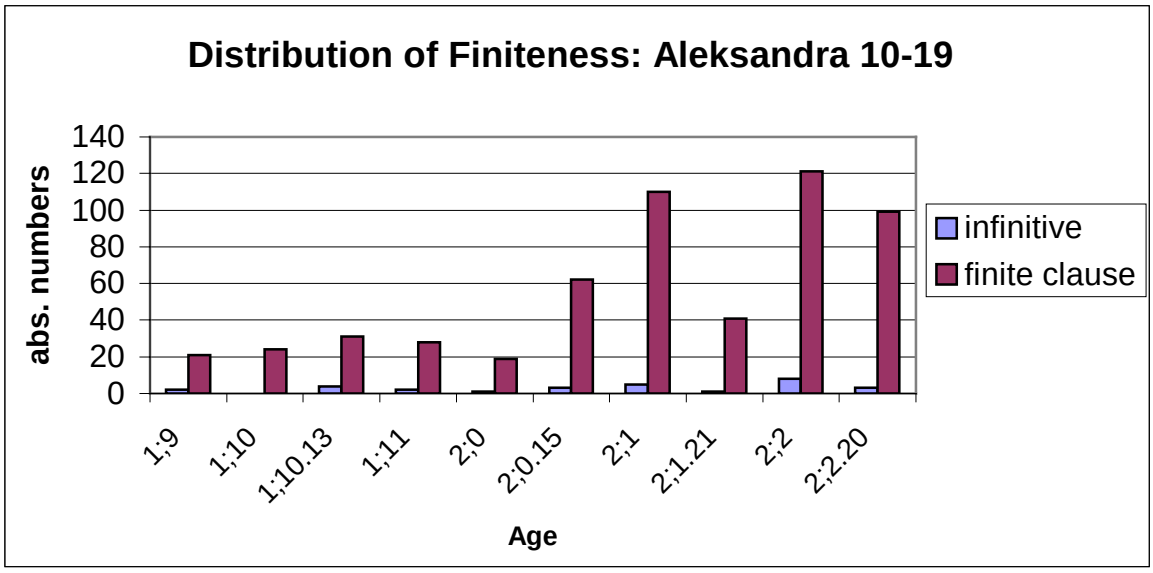


The first observation one might deduce from these tables and figures is a remarkably low frequency of non-finite verb forms in main clauses (main clause infinitives, MCI), as represented in the rightmost column of table 3 and 4. This rare occurrence of matrix infinitives can be observed in both corpora, but especially in the *Anna*-Corpus. This is particularly obvious in Figure 2b that shows the frequency of MCIs in percentage terms – a percentage approaching zero.

If we compare the two corpora, we can see different patterns in the two children, though. In the first four files of the *Dagmara*-Corpus (until the age of 2;04), frequency rates of main clause infinitives go up to 22,5 % (file 3, at the age of 2;04), but then there is a drastic decrease of non-finite forms that seems to remain constant after that point. However, even in those files (*Dagmara* 2-4) with a relatively high percentage of main clause infinitives, their frequency is still below the average of the OI languages. These MCIs will still undergo a qualitative analysis in section 6.1.3, where their interpretation will be studied.

Before drawing any hasty conclusions, however, we should rule out the possibility that the children analysed so far have already gone beyond the *Optional Infinitive Stage*, - in other words: that the crucial stage might start earlier than 2;0 in Polish child language. In order to check this possibility, we will analyse the earliest data from the *Aleksandra*-corpus that lend themselves to a quantitative analysis: the files from *Aleksandra* 10 (age 1;9) on. Figure 3a below illustrates the distribution of finiteness in the relevant files (*Aleks* 10-19) in absolute numbers, thereby completing the overall picture of the finiteness distinction in child language:

Figure 3a: Distribution of infinitives vs. finite verb utterances (*Aleksandra*-Corpus):



Since the entire *Aleksandra*-corpus is analysed according to the *data evaluation scheme* and the results of the counts of all files are given in the appendix, I will only present the corresponding diagrams here. The analysis is confined to the first ten quantifiable files of the *Aleksandra*-corpus, i.e. files *Aleks* 10-19. The actual figures and counts that Figure 3a and 3b are based on can be looked up in the tables with the data evaluation results in the appendix[1]. Figure 3b illustrates the distribution of finite vs. non-finite utterances in the relevant *Aleksandra*-files in per cent:

Figure 3b: Distribution of infinitives vs. finite verb utterances in % (*Aleksandra*-Corpus):

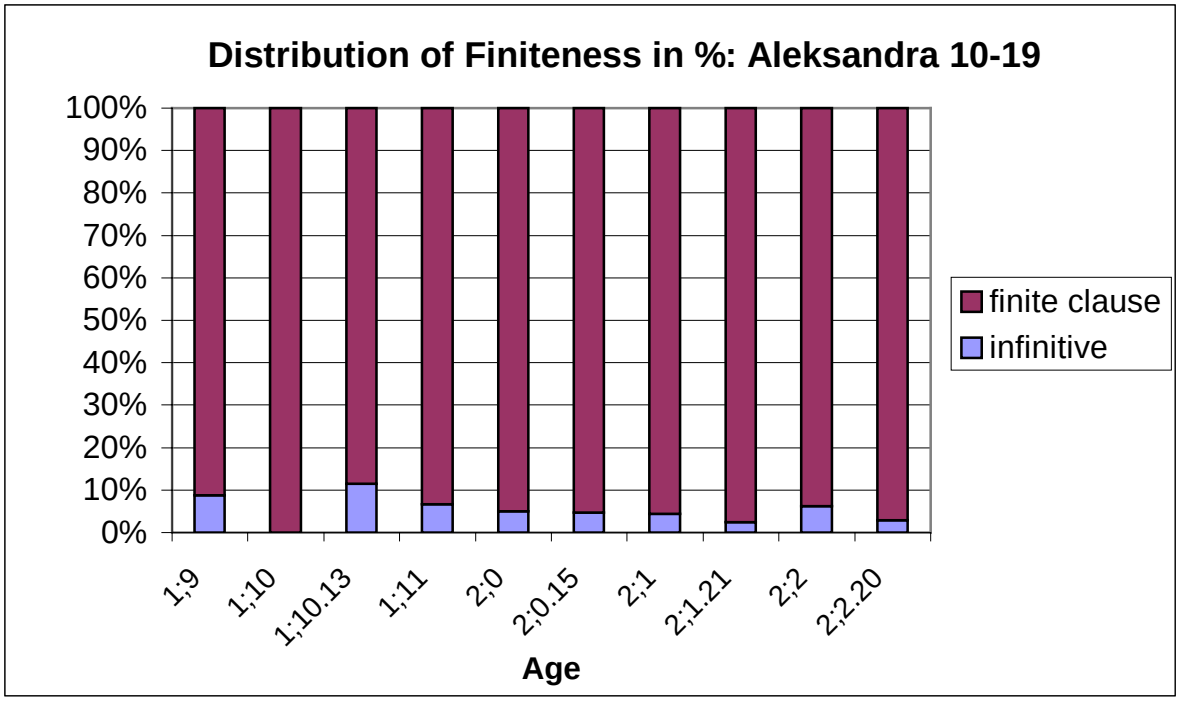


Figure 3b above clearly shows that the frequency of main clause infinitives is also very low in the earliest quantifiable data of the *Aleksandra*-Corpus. Only at one point (file 12, at the age of 1;10.13) can we observe a frequency rate of MCIs of 11.4% , - in all the other files, frequency rates of non-finite utterances are clearly below 10%.

One can tentatively conclude, therefore, that – across individual children and different stages of acquisition - , there seems to be no evidence for an *Optional Infinitive Stage* in Polish child

[1] The figures used for diagrams 3a and 3b represent the total number of finite/non-finite utterances of each file, including both null- and overt-subject utterances, as given in the first section of the scheme at Variable No. 1.

language data. The frequency rate of main clause infinitives in our data is significantly below the average of the typical OI-languages.

The issue of optionality, however, should not only be addressed in terms of *quantitative* analyses. In fact, an adequate approach to this subject matter should always include a *qualitative* analysis of the data as well. Formal criteria and frequency counts alone will not give us a satisfactory answer to the puzzle of this major case of optionality found in the area of verbal inflection. Is there really a coexistence of two verbal forms in young children's grammatical systems, one being the grammatical, finite verb form, and the other one being an "optional", non-finite form that is ungrammatical in the target language? Understanding the nature of these main clause infinitives definitely involves an approach to their meaning.

This issue leads us to the second part of research question Q1, which takes us from a quantitative to a qualitative analysis of the main clause infinitives of the three corpora.

6.1.3 The Interpretation of Main Clause Infinitives: Context Analyses

Let's begin by recalling research question Q1b, which is supposed to be addressed in this section:

Q1: b. As for the non-finite verb forms in main clauses (= "main-clause infinitives (MCI)"), - in what contexts do these MCIs occur and what is their intended meaning?

An answer to this question can only be given along the lines of a qualitative data analysis.
Before drawing any conclusions concerning the OI-status of Polish, two major aspects that seem to interact with this issue need to be considered: apart from the semantic aspect covered in Q1b, there might be a phonological issue that needs to be checked first. As mentioned above, the infinitive constitutes a highly marked form in the Polish verbal paradigm and, as such, it might cause phonological difficulties to very young children. We have to check, therefore, if the virtual absence of main clause infinitives in the earlier data of the *Anna*-Corpus especially might be due to pronunciation or articulation difficulties caused by the final

segment present in all infinitives, i.e. /-ć/, the prototypical infinitival suffix in the form of a palatalised affricate [tç]. One way to check this possibility is to re-examine the data and find out if the infinitive does in fact occur inside modal constructions, e.g. [auxiliary/modal verb + infinitive]. If it did, this would speak against a phonological explanation for the virtual absence of MCIs in the early data.

So therefore we checked the relevant files of the *Anna*-Corpus (*Anna* 1-11) for these modal constructions to see if the infinitive form occurred at all in the data. By the way, this question is less relevant for the *Dagmara*-Corpus, as these data already provide sufficient evidence for the occurrence of infinitives in the form of MCIs (with frequency rates for MCI going up to 22.5% at one point, see table 1). Therefore, I will only give the results for the *Anna*-Corpus here, but the *Aleksandra*-data basically show the same result (see the exact figures in the appendix).

The frequency rates for the occurrence of "complete" [aux./mod. + infinitive]-constructions (labelled **[mod+inf]** in the rightmost column of table 3) are given below for the *Anna*-Corpus (table 3), in absolute numbers and per cent:

Table 3: Occurrence of [Mod + Infinitive]-constructions: *Anna*-Corpus

File	Age	verb-utt.	Fin.V	/ %	Infin.	/ %	[Mod.+ Inf.]	/ %
1	1;11.23	82	71	86.6 %	4	4.9 %	3	3.7 %
2	2;0.13	70	67	95.7 %	0	0 %	3	4.3 %
3	2;01.3	76	73	96.1 %	1	1.3 %	2	2.6 %
4	2;01.25	92	82	89.1 %	3	3.3 %	7	7.6 %
5	2;02.26	98	85	86.7 %	7	7.2 %	5	5.1 %
6	2;03.20	89	86	96.6 %	1	1.1 %	4	4.5 %
7	2;04.13	165	157	95.1 %	3	1.8 %	23	13.9 %
8	2;05.4	72	68	94.4 %	1	1.4 %	5	7 %
9	2;05.23	77	74	96.1 %	2	2.6 %	5	6.5 %
10	2;06.15	124	118	95.1 %	2	1.6 %	14	11.3 %
11	2;07.6	111	107	96.4 %	2	1.8 %	20	18 %

If we look at the occurrence of "complete" [aux./mod. + infinitive]-constructions (both in absolute numbers and in per cent) and compare those figures to the occurrence of main clause infinitives, we can see sufficient evidence for the occurrence of infinitive forms in the early data. Since Anna seems to produce infinitive forms that are part of modal constructions from the very beginning and continues to use them more and more often throughout the recordings, we can safely rule out the possibility of phonological difficulties as an explanation for the virtual absence of MCIs in the *Anna*-Corpus.

One question that remains to be settled now concerns the interpretation of the main clause infinitives in the *Dagmara*- and the *Aleksandra*-Corpus (see Q1b). The best test case for this semantic question are the first four files of the *Dagmara*-Corpus (until the age of 2;04), when frequency rates of main clause infinitives go up to 22,5 % (file 3, at the age of 2;04). Afterwards, there is a drastic decrease of non-finite forms that seems to remain constant after that point. The main clause infinitives of the first files, however, still call for an explanation.

In principle, there are (at least) two possibilities. On the one hand, these main clause infinitives might be instances of *optional infinitives*, which would imply that they represent true alternations with finite verbs as in the OI-languages. Due to the overall low percentage of MCIs in the three corpora, however, this possibility does not seem to be very plausible.

On the other hand, one could imagine that these main clause infinitives have an altogether different meaning, i.e. an interpretation along the lines of Boser et al.'s (1992) alternative proposal. These authors try to account for the occurrence of main clause infinitives in different Germanic languages by postulating the *Null-Auxiliary Hypothesis* (see the discussion in chapter 3 for theoretical details, we will confine ourselves to the application of this approach here).

If we want to test this second explanation empirically, we have to approach the intended meaning of the MCIs with the help of a context analysis. That means that all instances of non-finite verbs in main clauses in the *Dagmara*-Corpus are analysed with respect to the **context** of the relevant utterances. The MCIs are classified as either **modal** (i.e. instances of [null-aux./null-modal + infinitive]-structures), **non-modal** or **unclear** cases. Non-modal contexts are standard instances of declarative contexts which represent neutral descriptions of events, whereas modal contexts involve the idea of volition, intention or futurity, usually referring to

non-real events. As for the *Aleksandra*-data, the *data evaluation scheme* (variable 8, see description in chapter 5) further differentiates between modal/volitional and elliptical/context-licensed cases.

To give an example of how we proceeded in this context analysis, let's have a look at a famous German example taken from Poeppel & Wexler (1993:5):

(1) Thorsten Ball haben.
Thorsten ball have(infin.)

If we want to decide on the intended meaning of (1), we have to look at the actual context of this utterance. Two readings are feasible here:

(1') 'Thorsten **has** a ball, i.e. he is holding a ball in his hands.' = declarative reading
(1") 'Thorsten **wants to/ will** have a ball.' = modal reading

It is only by context information (in this case as to whether the child is actually holding the ball in his hands while uttering (1) or whether he was more likely to express a wish or intention) that we are able to disambiguate between (1') and (1"). Only in the first case could we safely assume that the child utterance is an instance of an *Optional Infinitive*, i.e. a non-finite utterance with a declarative meaning.
This is where the importance of reliable data documentation and annotation becomes evident (see the methodological discussion in section 5.1). Without contextual or situational information, we would be unable to determine the intended meaning of the data.

It should be noted, however, that any attempt at assessing the "intended reading" of a child's utterance remains problematic. Even with the help of carefully documented data and detailed context information it is sometimes impossible to assign an (unambiguous) interpretation to a given utterance (e.g. "unclear cases" in table 4).

The results of this context analysis for the *Dagmara*-Corpus are given in Table 4 below.

Table 4: Context-Analysis of non-finite matrix clauses: *Dagmara*-Corpus

File / Age	Non-fin. utt. (n./%)	Modal (n./%)	Non-modal (n./%)	Unclear (n./%)
1 / 2;2.25	21 / 10.2%	16 / **76.2 %**	2 / 9.3 %	3 / 14.3%
2 / 2;3.14	22 / 15.5%	18 / **81.1 %**	1 / 4.6 %	3 / 13.6%
3 / 2;4.3	41 / 22.5%	30 / **73.2 %**	3 / 7.3 %	8 / 19.5%
4 / 2;4.20	33 / 16.8 %	25 / **75.7 %**	3 / 9.1 %	5 / 15.2%
5 / 2;5.13	9 / 3.7 %	7 / 77.8 %	2 / 22.2 %	0 / 0 %
6 / 2;6.1	3 / 1.4 %	3 / 100 %	0 / 0 %	0 / 0 %
7 / 2;7.3	1 / 0.5 %	1 / 100 %	0 / 0 %	0 / 0
8 / 2;7.21	1 / 0.5 %	1 / 100 %	0 / 0 %	0 / 0 %

The results of the counts for the *Aleksandra*-data can be directly read from the evaluation tables (variable 8) in the appendix. Therefore, they won't be repeated here. To summarize the findings briefly, out of all the main clause infinitives in the relevant files *Aleks* 10-19 (age 1;9-2;2) – which are 29 in total – there are only two cases where the interpretation is unclear. 21 instances of MCIs clearly show a modal or volitional meaning, and 6 of them are elliptical or context-licensed (often with future meaning). Crucially, there are *no* instances of main clause infinitives with a *non*-modal or *non*-elliptical reading.

On the whole, we can conclude from the context analysis of both corpora that non-finite verb forms in main clauses mainly occur in modal or elliptical contexts where they are pragmatically licensed. What do these utterances actually look like?
Let's illustrate this by the following examples which were taken from the *Anna*- Corpus (2) and the *Dagmara*-Corpus (3): (M= Mother, A= Anna, D= Dagmara)

Modal contexts: Situation: Playing doctor. Anna gives the thermometer to her mother.

(2) M: I co z tym termometrem? Co mam zrobić?
'What am I to do with this thermometer?' -

A: Już zmierzy**ć.** [*Anna* 5, 11 (2;2)]
Already measure- INF
'You must/should take your temperature already.'

Situation: Dagmara points to a book, saying:

(3) D: Z mamą pooglądać tolipanów (tulipany) [*Dag* 3,5 (2;4)][2]
with mummy look-at- INF tulips
'I want/would like to look at the tulips with mummy.'

Throughout the two corpora, we only find very few examples of true alternations between finite and non-finite forms for declarative (= non-modal) contexts. One of the rare instances of a real [+/- finite] alternation, which appears to be parallel to the examples quoted for OI-languages, is given in (4). In this utterance, the child seems to play with optionality, since she exploits the (early child grammar) option of choosing either a finite or a non-finite verb form for declarative/non-modal contexts. The two utterances in (4) immediately follow one another during the recording. In the target system, of course, the main clause infinitive in (4) is not a grammatical option for declarative contexts.

(4) Situation: We are cooking milk for Dagmara's teddy.

D: Mleczko mu **ugotujemy**. [*Dag* 1,171 (2;2)]
Milk (for) him cook-1.PL.PRS
'We are cooking milk for him.'

Niespodziankę jakąś mu **ugotować.** [*Dag* 1, 172 (2;2)]
Surprise(Akk) some him cook-INF
'We are cooking some sort of surprise for him.'

Considering the fact that utterances such as (4) represent rare examples of true alternations between finite and non-finite verbs, we do not have any evidence for the OI phenomenon in Polish (at least on the basis of the three corpora studied here).

6.1.4 Summary

Let's briefly summarize what we have seen so far. The analysis of the distribution of finiteness (6.1.2) has shown that in all three corpora, there seems to be no evidence for a genuine *Optional Infinitive Stage* in Polish. The data were analysed both quantitatively (section 6.1.2) and qualitatively (6.1.3). As for the latter, we analysed the main clause

[2] The reference for the child data quoted from my databasis always follow the same structure: File name, File number, utterance number, and the child's age (given in brackets, according to the formula (years;months.days)).

infinitives in the two corpora where MCIs occurred most frequently – the *Dagmara*- and the *Anna*-Corpus – with respect to their intended meaning. The context analyses for both corpora point to two observations:

On the one hand, the control counts (see the rightmost column of table 3) show that infinitival forms do occur in "complete" [auxiliary/modal + infinitive]-structures in both corpora, even in the *Anna*-Corpus that shows an extremely low percentage of MCIs. This fact provides counterevidence to the idea that infinitives may be absent due to pronunciation difficulties (most infinitival forms end in the palatalised version of the affricate [tç] which is indeed difficult to pronounce). The data, however, clearly show that infinitives do occur as part of "complete" [modal + infinitive]-constructions and that they even do so with considerable frequency (again, especially in the *Dagmara* data).

Secondly, and more crucially, the data seem to indicate that infinitives in matrix clauses mainly occur in modal contexts. These root infinitives are often used in utterances that represent answers to a question containing a modal or auxiliary verb (cf. (2)). That means that examples such as (2) and (3) tend to be pragmatically licensed and represent, therefore, possible grammatical structures of the target language as well. The majority of main clause infinitives occurring in modal contexts involve volitional or future tense utterances in which the modal or auxiliary appears to have been dropped.

Before drawing any general conclusions, however, another possibility needs to be considered. We should look at an even earlier stage of language development that is usually hard to capture, i.e. the one- and two-word-stage. What do the very first verb forms a child utters look like?

6.2 A Glance at the Earliest Stage of Acquisition

6.2.1 Dealing with Early Speech Data: *Aleksandra*-Corpus

Since the *Dagmara*- and *Anna*-data do not cover the earliest developmental stages (i.e. the very beginning of child language production), it might be the case that we simply "missed" an earlier period of use of non-finite matrix verbs. In order to check this possibility, we extended the database, that originally consisted of the *Dagmara*- and *Anna*-Corpus only, to the *Aleksandra*-corpus which covers the whole range of acquisitional stages, starting from the utterance of single words up to and including multi-word utterances.

Dealing with early speech data poses a lot of problems, however. One side of the problem is due to the scarceness of early child language data. As very young children are often quite hesitant or even reluctant to talk, it might turn out to be rather difficult to elicit any spontaneous utterances, - especially the kind of data that would still meet the *spontaneous speech* criteria discussed in chapter 5.

Being aware of these difficulties, the mother would break up the recordings into smaller sessions in the earliest phase of the study, depending on how communicative the child was. During these early stages, the mother also recorded a lot of spontaneous situations, i.e. whenever she felt the child was willing to communicate. Thanks to her sensitive way of handling the recordings, we are provided with a reliable and revealing documentation of the emergence of the very first verbs in Aleksandra's language development.

The other side of the problem with early child data has to do with the interpretation of early utterances. As early utterances merely consist of single words in isolation, their meaning is often difficult to assess (and mostly impossible without context information). Therefore, we added contextual information to the individual child utterances in order to be able to assess the intended meaning. The recordings were done by the child's mother who was also part of the transliteration team, and she was most likely to infer the intended meaning, even in some critical cases where the interpretation was difficult. For any critical child utterance, I made a note of both linguistic and non-linguistic aspects that seemed to be relevant to the interpretation. As for the linguistic aspects, I would list grammatical features of the child utterances and – in case of ambiguous or deviant structures – the corresponding correct

sentence in the target language. Non-linguistic notes would typically include information on the situational context of the utterance as well as a brief description of the child's actions while speaking.
Of course, there is still no guarantee that one obtains *the* one and only interpretation, but at least one can approach it as far as possible.

As a result, the *Aleksandra* data are likely to provide a very reliable basis for completing the overall picture of acquisition since they allow us to trace the course of development back to the very beginning, i.e. to the utterance of the first words. Thanks to the regularity of the recordings, it is possible to follow every single step throughout the acquisition process.

6.2.2 The Nature of Earliest Verb Utterances (One-and Two-Word Stage)

We will now go back to the very beginning of language production in the *Aleksandra*-Corpus and have a look at the nature of the earliest (verb) utterances at the one- and two-word stage (files 1-9).
Table 5 gives a survey of the data we will be concerned with in this section:

Table 5: *Aleksandra*-Corpus: One-and Two-Word Stage (Aleks1-9)

	File	Age	Length
One-	1	1;04.15	30 min.
Word-	2	1;05.5	2 x 15 min.
Stage	3	1;05.14	15 min.
Transition	4	1;06.8	15 min.
to Two-	5	1;06.10	15 min.
Word-	6	1;06.14	15 min.
Stage	7	1;07.4	20 min.
Two-Word-	8	1;07.15	20 min.
Stage	9	1;08.7	20 min.

Since these first files covering data from the one- and two-word stage (i.e. up to file 9) do not lend themselves to a quantitative analysis, they are analysed from a more qualitatively orientated point of view, along the lines of our second research question:

Q2. What do the *earliest* stages of acquisition reveal about the clausal architecture of early child grammar?

More specifically:

a. What do the earliest verb utterances at the one- and two-word stage (*Aleks 1-9)* look like?

b. What kind of acquistional *mechanisms and strategies* can be deduced from the earliest part of the *Aleksandra*-data?

The results regarding Q2 a can be summarized as follows:
During the first two recordings of Aleksandra (at the age of 1,4 to 1;5), there are hardly any verbal forms to be found. Most utterances consist of names or nouns without any modifying material.
When the first verbal utterances appear, i.e. starting from file 3/4 at the age of 1;05/1;06, only two forms occur: 3rd person singular present tense and the imperative singular. The imperative seems to be the most frequent verbal form in these very early recordings. What is even more crucial, these early imperatives often tend to occur in declarative or descriptive contexts that require finite forms in the target grammar. Some infinitival constructions also refer to an ongoing activity, which can be demonstrated on the basis of the relevant context information in the transcripts. Two representative examples of this use are given in (5) and (6) below:

(5) While sitting down, Aleksandra says:

Siada**j**. [*Aleks* 2, 7 (1;05)]
'sit down**(imp)**'.
'Sit down!

(6) While giving a pen to her grandmother, she says:

Da**j.**
'give(**imp**)'.
'Give!' [*Aleks* 3, 14 (1;05)]

Even in contexts where the infinitive is required in the target language (for example in answers to questions containing [modal verb + infinitive]-constructions), Aleksandra still uses imperatives:

(7) M: Co Michał będzie robi**ć?** [mother's input]
What Michael will do-INF
'What will Michal do?' -

A: Cita**j**
read-**IMP**
'Read!' [*Aleks* 5, 9 (1;06)]

We can safely conclude, therefore, that even at this early stage in the acquisition of Polish, the infinitive does not occur in the way it has been claimed for many other (OI-)languages.

Instead, the imperative seems to play a special role in these earliest stages of acquisition, until the end of the two-word stage roughly, i.e. until the age of 1;8. Crucially, the imperative is used in declarative contexts, and is mostly associated with descriptive meaning (see example (5) and (6) above). We may tentatively suggest that the imperative seems to represent a kind of default form for the acquisition of verbal inflection in Polish in these earliest stages. Later on, i.e. from the age of 1;9 on, this option of imperatives used in declarative contexts ("*Descriptive Imperatives")* seems to be gone.

Before looking at this phenomenon in detail and discussing it from a cross-linguistic perspective, let's look at the occurrence of infinitives. Again, we have to consider the possibility of phonological difficulties being responsible for the absence of infinitives in the earliest files.

The first infinitival forms in the *Aleksandra*-Corpus occur in file 9 at the age of 1;8, i.e. at the transitional phase between the two-word- and the multiword stage. These first non-finite forms, however, do not occur as main clause / *optional* infinitives, but in combination with auxiliaries to form the future tense as in (8) or together with modal verbs to form a [modal + infinitive]-structure. This first instance of an infinitive in the *Aleksandra*-Corpus is given in (8):

(8) Tutaj będzie Zizi mieszkać. [*Aleks* 9, 17 (1;08)]
Here will- 3.SG.FUT Zizi live-INF
'Zizi will live here.'

Early utterances like (8) again provide counterevidence to the idea that infinitives may be absent due to pronunciation difficulties.
Instead, data like (8) show that infinitival forms do occur in "complete" [auxiliary/modal + infinitive]-structures, and – on top of that – even *before* the occurrence of the first main clause infinitives.

6.2.3 The Development of Subject-Verb Agreement

The surprising finding of descriptive imperatives might reveal some of the acquisition mechanisms that young children use during the formation of the grammatical system of their target language. Let's approach this issue by raising some questions that are linked to it:

- Why do children initially use imperatives (a very frequent form in their input) in declarative contexts?
- Is this option due to a deficit in morphological or syntactic knowledge or does it rather show difficulties with the process of *accessing* morphological knowledge?
- On the assumption that there is in fact some deficit in knowledge, can we find any independent evidence for lack of morphological or syntactic knowledge in other areas of grammar in the data?
- Are there any cross-linguistic findings that show the same or a similar phenomenon?

These questions will lead us through the argumentation in this and the next section.

Let's look at the question of a possible syntactic or morphological deficit first. If we want to assume that children's earliest grammatical system is missing a crucial item of knowledge, we would expect evidence for lack of knowledge in other areas of grammar as well. Since the phase in question is very short and the data scarce, it is not so simple to test this claim empirically. A suitable test case, however, is the use of verbal inflection in child language.

As for the *Aleksandra*-data, we can safely claim that whenever a finite form appears in the data, this form is used correctly with respect to verb morphology and subject-verb agreement. Throughout the *Aleksandra*-Corpus, there are hardly any instances of agreement errors (see also Guasti (1993/4) for similar findings for Italian). Let's illustrate this claim by looking at

two representative utterances from the first file that was analysed according to the data evaluation scheme, file 10 (1;9):

(9) Duda mieszkała w tym domtu. [domku]
Duda live-3.SG-PAST-FEM in this house.
'Duda lived in this house.' [*Aleks* 10, 13 (1;9)]

(10) pszytyjemy tym. [przykryjemy]
Cover-1. PL-PRES with this
'We cover (it) with this.' [*Aleks* 10, 4 (1;9)]

These two utterances that were produced at the age of 1;9 provide striking evidence for differentiated morphosyntactic knowledge at a very early stage of acquisition. One item of knowledge is reflected in the correct morphological expression of agreement. In the case of (10), this knowledge is even expressed in the use of plural inflection (1st-person plural), although plural person markers are usually said to be delayed (probably due to a more general delay in the use of plurality). On top of that, (9) shows correct use of tense markers (in this case past tense) and even gender agreement markers (i.e. –a).

As far as errors are concerned, we can only observe some errors of *omission* rather than incorrect forms. In fact, there is only one *recurrent* error, the omission of the (invariable) reflexive pronoun się (e.g. **się** bać 'to be afraid'), which is ungrammatical in the target language. These errors of omission are noted down in the data evaluation tables (appendix) under Variable 10, value 6 "missing pronoun", including examples (next to the absolute numbers). The omission of the reflexive się seems to represent a true option, involving an alternation of [+/- reflexive], since the child uses structures with overt and dropped reflexives side by side, sometimes in the same discourse, as the following dialogue shows:

(11) A: A ty **się** nie boisz?
And you refl. not be-afraid-2SG [*Aleks* 18, 146-148, (2;2)]
'And you are not afraid?'

M: Ja trochę **się** boję.
I a little refl. be-afraid-1SG
'I am a little afraid.'

A: Trochę e boisz? A dużo e nie boisz?
A little be-afraid-2SG and much not be-afraid-2SG
'Are you a little afraid?' 'But you are not much afraid?'

In conclusion, we can say that neither in the earlier nor in the later files is there any evidence for a significant percentage of agreement errors (see the data evaluation tables at Variable 4-5 in the appendix, where all the different inflections are listed with concrete figures regarding their frequency).

In addition to that, Aleksandra sometimes uses the same verb (e.g. śpiewać 'to sing') in a variety of morphosyntactic environments, e.g. as a *Descriptive Imperative*, a finite verb form, and a bit later also as en elliptical infinitive. This finding, together with her correct use of the imperative and her correct finite forms, indicates a very early ability to distinguish appropriately between finite and non-finite verbs.

Data from other languages (Italian, for instance, see Guasti (1993/4)) tend to confirm this result. We will now turn to some cross-linguistic evidence that might shed additional light on the acquisitional mechanisms to be postulated for children's early grammatical systems.

6.2.4 Mechanisms in the Early Formation of Grammar: Cross-Linguistic Evidence for *Surrogate Verb Forms*

Another particularly interesting language to look at is Russian, a language that, though typologically similar, shows at least two crucial differences, compared to Polish: it is a non-pro-drop (or non-null-subject-) language, and it does show an *Optional Infinitive Stage.* The Russian children studied in previous research were shown to use finite verb morphology correctly from early on, even though they pass through an *Optional Infinitive Stage* between the ages 1;10 and 2;6 (Bar-Shalom & Snyder (1996, 1998)). In a recent study, Bar-Shalom & Snyder (2001) present new evidence from the *Svetlana*-Corpus that covers an even earlier period (from the age of 1;8 on) than the other Russian child data studied so far.

By referring to my earlier papers (Klepper-Schudlo (1996, 1999)), the above authors discover a very interesting new aspect of these earliest acquisitional stages of Russian that appears to resemble my findings for Polish: the early use of imperatives with a descriptive meaning ("*Descriptive Imperatives").*The following examples (taken from Bar-Shalom & Snyder (2001:95)) illustrate the descriptive use of imperatives in the *Svetlana*-data:

(12) Child: Oden' <zakolku>
put-on-**IMP** pin [Svetlana, 1;8]

(13) Mother: A kto eto na kresle sidit?
And who in armchair sits?
'And who is sitting in the armchair?'

Child: Sazhaj.
Seat-**IMP** [Svetlana, 1;8]

According to the authors, the context for (12) was that Svetlana was trying to put a pin in her doll's hair. Her utterance thus described her own, on-going action. This interpretation is even reinforced by the fact that her mother mistakenly believed that Svetlana was uttering a true imperative, i.e. that was asking for help in inserting the pin. When her mother tried to help, however, Svetlana grabbed the doll away.
In (13), Svetlana and her mother were discussing a picture of some children sitting in chairs. Again, the use of the imperative refers to an on-going action.

The Polish and Russian results are supported by similar findings in other languages, notably pro-drop / null-subject languages. In Croatian child language data, for instance, there seems to be evidence for an overgeneralized use of both imperatives and 3rd person-singular forms for various verbal categories in the earliest stages. Katičić (1997:51ff. & 106ff.) reports the following examples from her *Antonija*-Corpus to illustrate this phenomenon:

(14) cavi [=stavi]
here put-**IMP** [*Antonija* 9 (1;07)]

(15) Vidi ja
see-**3. SG** I [*Antonija* 9 (1;07)]

The use of the 1st-person personal pronoun '*I*' in (15) shows that the verb which is marked as 3rd-person singular by a null-morpheme of the present tense stem (Katičić (1997:55ff.) cannot really be interpreted as a genuine 3rd-person singular form. Katičić quotes a few more examples from the *Antonija*-Corpus showing that this form is more likely to represent an "overgeneralized form" that is used for different inflectional categories. As the 3rd-person singular formation corresponds to the present tense stem null-morpheme, its deviant use in the early Croatian data might be interpreted as an optional use of the least marked form in the verbal paradigm.

Similar findings are reported in Dressler & Makovec-Cerne (1995) for the early stages in the acquisition of Slovenian, when the child initially uses imperative forms in declarative contexts, parallel to my findings for Polish.

Considering all these results from different languages, how could the phenomenon of *Descriptive Imperatives* and other "overgeneralized" forms be captured in a unified account?

Bar-Shalom and Snyder (2001) suggest that both Polish and Russian *Descriptive Imperatives* could result from the children's misanalysis of the imperative as a bare stem. On this view, the Bare Stem Parameter (cf. Hyams (1986)) would sometimes remain unset, at early ages, for Russian and Polish children. Hence, the Russian and Polish children would initially mistake the imperative for a bare stem, and would mistakenly believe that bare stems are morphologically permissible as surface forms, resulting in *Descriptive Imperatives.*

Any account of this phenomenon, however, has to come to terms with the fact that children acquiring different rich-agreement languages (including Italian, Russian and Polish) were shown to be able to distinguish between finite and non-finite verbs, and, thereby, were shown to possess morphological and syntactic knowledge from the earliest stages of language acquisition on. What might be missing initially, however, is a fast, automated process of *accessing* that knowledge.
By differentiating between very young children's nearly error-free morphological knowledge and the less reliable *implementation* of that knowledge, we can reconcile two apparently conflicting findings and gain a unified account. Therefore, I would take these *Descriptive Imperatives* to be a natural "default" form that is used when the child fails to retrieve the relevant morphological realization. During the earliest stages of language acquisition, the child would use the imperative as a "surrogate" verb form whenever the features inserted in the inflectional system cannot otherwise be expressed. Their semantic interpretation, however, remains unchanged, even when their morphosyntactic features have been neutralized.

This issue will be picked up again in the general conclusion in the next chapter, when I will try to link this account of *Descriptive Imperatives* to the *Optional Infinitive* phenomenon. I will extend the cross-linguistic perspective even further by looking at data from a few other languages that show evidence for *surrogate verb forms,* such as Austrian German, Greek and Inuktitut.

6.3. The realization of subjects in the *Aleksandra*-Corpus

In the previous section, I have argued in favour of children knowing the morphosyntactic properties of verbs. Finite verbs are usually associated with agreement and tense features. The fact that the Polish children use agreement morphemes correctly whenever they produce a finite utterance, seems to indicate that early child grammar contains the relevant agreement features.
Let's look at a related issue, the realization of subjects in the *Aleksandra*-Corpus and the distribution of null-subjects vs. overt subjects in the data.

This topic again is related to the issue of o*ptionality* in early child grammar. In fact, dealing with null-subjects means entering into an area that is at the forefront of current research. One of the most salient properties of early language is the apparent optionality of subjects. All children pass through a stage in which they frequently omit subjects, even (or should I say: *particularly)* if the target language is *not* a null-subject language. In recent publications, many studies demonstrate the omission of subjects in early acquisitional stages in languages that do not permit null-subjects. In the present study, we will approach the phenomenon from the other side and see how the distribution of null- vs. overt subjects looks like in Polish (both child and adult language), a language that is traditionally classified as a null-subject / pro-drop language. We will return to this cross-linguistic issue in section 6.3.4, where we will confront our findings for Polish with results and data from other languages.

6.3.1 The distribution of null- vs. overt subjects

In this section, we will address the third research question (Q3) that is repeated here:

Q3. How does the realization of subjects develop in the *Aleksandra*-data?

Concretely:

a. What does the distribution of null- vs. overt subjects look like?

b. What *types* of subjects can be identified and how are these subjects types (notably NP- vs. pronominal subjects) distributed?

As discussed in chapter 4, Polish is traditionally classified as a null-subject language, which means that the subject is usually omitted (unless it is emphasized). Therefore, one would not

expect a high frequency of overt subject utterances in Polish child language. As a matter of fact, however, the Polish child data point to the opposite direction: there is a decisive percentage of utterances with an overt subject. This is illustrated in the following tables: Tables 6a-c below illustrate the distribution of null-subjects vs. overt subjects in both absolute numbers and per cent for every single file of the *Aleksandra*-Corpus. Table 6a covers files 10-18, Table 6b files 19-27, and Table 6c the last files, file 28-37:

Table 6a: ***Aleksandra*** files 10-18 (1;9-2;2): **Null-subj.- vs. overt subject utterances**

***ALEKS* files: 1;9 - 2;2**	**10 1;09.14**	**11 1;10.0**	**12 1;10.13**	**13 1;11.9**	**14 2;0.01**	**15 2;0.15**	**16 2;0.29**	**17 2;01.21**	**18 2;02.3**	**total 1;9 - 2;2**
# null subj.utt.	14	16	21	24	15	32	80	35	73	310
in %	66,6	66,6	67,7	85,7	78,9	51,6	72,7	85,4	60,3	67,8% av
# overt subj.	7	8	10	4	4	30	30	6	48	147
in %	33,3	33,3	32,3	14,3	21,1	48,4	27,3	14,6	39,7	32.2% av
Fin.utt. total	21	24	31	28	19	62	110	41	121	457

Table 6b: ***Aleksandra*** files 19-27 (2;2-2;6): **Null-subj.- vs. overt subject utterances**

***ALEKS* files: 2;2 - 2;6**	**19 2;02.20**	**20 2;03.4**	**21 2;03.18**	**22 2;04.4**	**23 2;04.18**	**24 2;05.1**	**25 2;05.15**	**26 2;06.0**	**27 2;06.16**	**total 2;2 - 2;6**
# null subj.utt.	56	72	69	50	53	51	47	75	81	554
in %	56,6	53,3	62,7	56,8	46,1	54,8	65,3	64,1	60,9	57,6% av
# overt subj.	43	63	41	38	62	42	25	42	52	408
in %	43,4	46,7	37,3	43,2	53,9	45,2	34,7	35,9	39,1	42.4% av
Fin.utt. total	99	135	110	88	115	93	72	117	133	962

Table 6c: ***Aleksandra*** files 28-37 (2;7-3;3): **Null-subj.- vs. overt subject utt.**

***ALEKS* files: 2;7- 3;3**	**28 2;07.0**	**29 2;07.14**	**30 2;08.0**	**31 2;08.15**	**32 2;09.7**	**33 2;10.1**	**34 2;11.7**	**35 3;0.19**	**36 3;01.24**	**37 3;03.2**	**total 2;7- 3;3**
# null subj.	74	87	68	60	64	65	63	72	94	112	759
in %	54	59,6	55,3	56,6	52,0	59,6	57,3	69,9	61	63,3	58.9% av
# overt subj.	63	59	55	46	59	44	47	31	60	65	529
in %	46	40,4	44,7	43,4	48,0	40,4	42,7	30,1	39	36,7	41.1% av
Fin.utt. total	137	146	123	106	123	109	110	103	154	177	1288

If we compare the three tables that represent three subsets of the *Aleksandra*-Corpus, we can see that from file 19 (2;02) on, the percentage of overt subjects is consistently above 40% on average. The highest frequency rate seem to occur between the age of 2;02 and 2;11, when the percentage of overt subjects goes up to 53.9% at one point (*Aleks* 23, at the age of 2;04).
Table 7 below gives an overview of the entire *Aleksandra*-Corpus, summarizing the results of tables 6a-c:

Table 7: **Overview: Null subj.** vs. **overt subj. utt.** (Aleksandra corpus: **F10-37**)

***ALEKS* files:** **Age:**	**F. 10-18** **(1;9-2;2)**	**F. 19-27** **(2;2-2;6)**	**F. 28-37** **(2;7-3;3)**	**Total** **(1;9-3;3)**
Null subject utt.	310 67,8%	554 57,6%	759 58,9%	1623 59,9% av.
Overt subj. utt.	147 32,2%	408 42,4%	529 41,1%	1084 40,1% av.
Finite vb. utt. total	457	962	1288	2707

The decisive percentage of overt subjects is a surprising finding in view of the null-subject status of Polish. In one file (*Aleks* 23, 2;04) the number of utterances with an overt subject even exceeds the number of null-subjects. Crucially, at least 90% of these overt subjects are used *without* emphasis, i.e. with neutral intonation. This seems to indicate that the child uses both options – i.e. null-and overt subjects - in the same environments. From the age of 3;0 on, the percentage of overt subjects seems to decline again.

One obvious question that comes to mind here should be checked independently: what does the distribution look like for the target or adult system? Therefore, we analysed the mother's speech in three representative files with respect to the distribution of null- vs. overt subjects, in order to compare the frequency rates found in the child data to the distribution in adult speech. Since we also checked the type of subject involved in the mother's speech, the results will be presented in the next section.

Figure 4a below illustrates the development of the variable [+/- overt subject] between the age of 2;02 and 3;03. This diagram is based on absolute numbers, whereas Figure 4b illustrates the same development and the same period in per cent.

Figure 4a: Distribution of Null- vs. Overt Subject Utterances in absolute numbers: (*Aleksandra* 18-37)

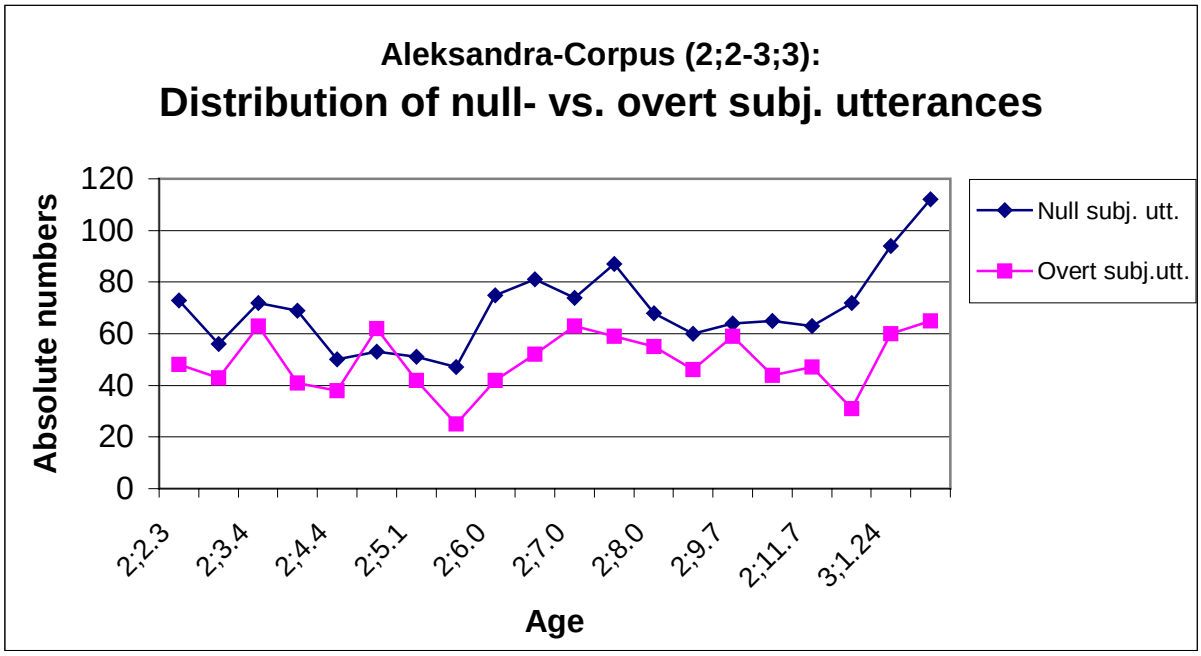
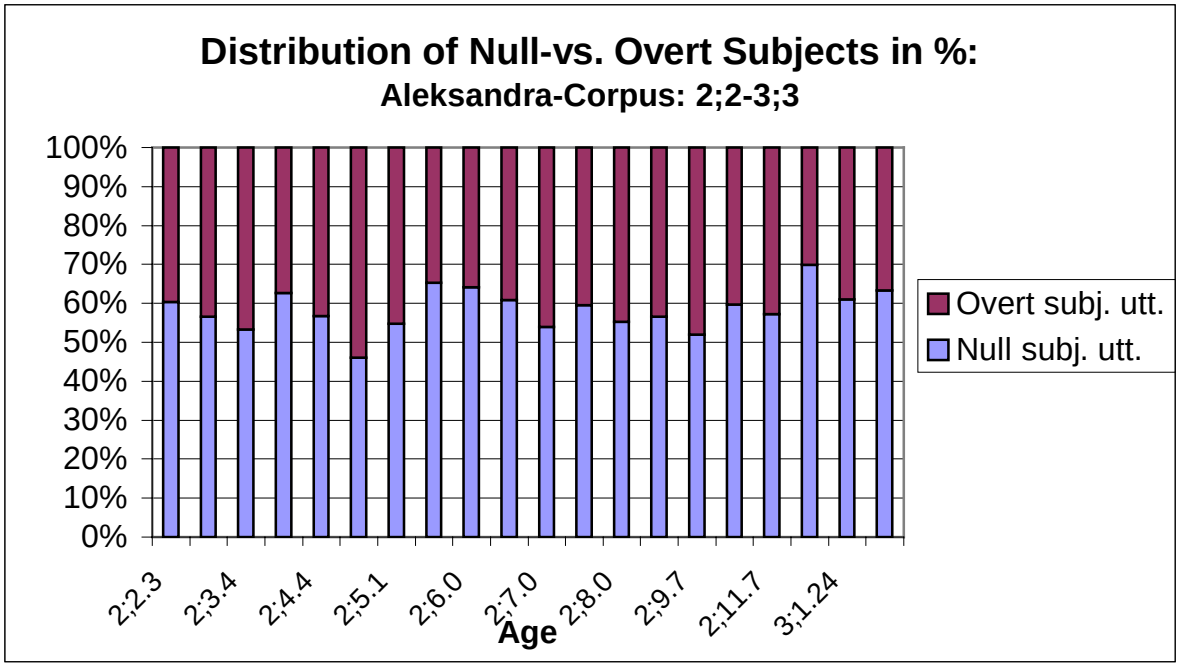


Figure 4a illustrates very clearly that null subjects and overt subjects are almost evenly distributed for a period of one year approximately. The same development is shown in per cent in Figure 4b:

Figure 4b: Distribution of Null- vs. Overt Subject Utterances in %: (*Aleksandra* 18-37)



6.3.2 Subject Type: NP vs. Pronominal Subject

Given the surprisingly high percentage of child utterances with overt subjects, it seems worthwhile to investigate the *type* of subjects involved (cf. Q 3b). As mentioned in chapter 5, the data evaluation scheme contains one variable (Variable 9) that specifies an overt subject with respect to its lexical content (cf. the tables with the data evaluation results in the appendix, see Variable 9: ”Lexical content overt subject”). We distinguish three types of overt subjects: Noun / NP, personal pronoun or other pronouns (e.g. demonstrative pronouns). As we are particularly interested in the general issue of the clausal architecture and the presence of functional categories in early child grammar, the occurrence of personal (subject) pronouns is of special interest to us.

We only consider the files from *Aleks* 18 (2;02) on, since these were the most productive files of the *Aleksandra*-Corpus with respect to the occurrence of overt subjects.

The first table below (Table 8 on the next page) gives an overview of the entire corpus, listing the individual distribution of subject types for every individual file. Subject pronouns and other pronouns (i.e. demonstrative pronouns) are first listed in two separate columns that are then taken together (“pronominal subject total”). This allows us to compare the frequency of both pronominal subjects in total and each pronoun category separately to the frequency of N/NP-subjects. All figures are given in absolute numbers as well as in per cent. For comparison, the rightmost column gives the total number of overt subject utterances as well as their percentage out of the set of finite utterances. Table 8 (next page) lists the counts for each file separately in order to give an exact and realistic picture of the overall development. For a better overview, Table 9 summarizes these findings in a more condensed form, grouping six files together at a time.

Table 8: ***Aleksandra*** files 18-37: **Subject type: NP vs. pronominal subject**
(Age: 2;2 - 3;3)

ALEKS files	N/NP subject (# / %)	subj. pronoun (# / %)	demonst. pron (# / %)	**pronom.subj. total** (# / %)	overt subj. total % (out of fin.utt)
F18 (2;2.3)	34 70,8%	9 18,8%	5 10,4%	14 29,2%	48 (= 39,7%)
F19 (2;2.20)	24 55,8%	14 32,6%	5 11,6%	19 44,2%	43 (= 43,4%)
F20 (2;3.4)	24 38,1%	31 49,2%	8 12,7%	39 61,9%	63 (=46,7%)
F21 (2;3.18)	26 63,4%	13 31,7%	2 4,9%	15 36,6%	41 (=37,3%)
F22 (2;4.4)	19 50%	10 26,3%	9 23,7%	19 50%	38 (= 43,2%)
F23 (2;4.18)	35 56,5%	17 27,4%	10 16,1%	27 43,5%	62 (= 53,9%)
F24 (2;5.1)	8 19,0 %	24 57,1%	10 23,8%	34 81,0%	42 (= 45,2%)
F25 (2;5.15)	13 52%	6 24%	6 24%	12 48%	25 (= 34,7%)
F26 (2;6.0)	24 57,1%	14 33,3%	4 9,5%	18 42,9%	42 (= 35,9%)
F27 (2;6.16)	26 50%	18 34,6%	8 15,4%	26 50%	52 (= 39,1%)
F28 (2;7.0)	25 39,7%	32 50,8%	6 9,5%	38 60,3%	63 (= 46,0 %)
F29 (2;7.14)	30 50,8%	21 35,6%	8 13,6%	29 49,2%	59 (= 40,4%)
F30 (2;8.0)	28 50,9%	24 43,6%	3 5,5%	27 49,1%	55 (= 44,7%)
F31 (2;8.15)	16 34,8%	16 34,8%	14 30,4%	30 65,2%	46 (= 43,4%)
F32 (2;9.7)	28 47,5%	22 37,3%	9 15,2%	31 52,5%	59 (= 48%)
F33 (2;10.1)	13 29,5%	20 45,5%	11 25%	31 70,5%	44 (= 40,4%)
F34 (2;11.7)	16 34%	23 48,9%	8 17,1%	31 66%	47 (= 42,7%)
F35 (3;0.19)	14 45,2%	17 54,8%	0 0%	17 54,8%	31 (= 30,1%)
F36 (3;1.24)	25 41,7%	34 56,7%	1 1,6%	35 58,3%	60 (= 39%)
F37 (3;3.2)	27 41,5%	32 49,2%	6 9,2%	38 58,5%	65 (= 36,7%)

The following table summarizes the findings displayed above in a more condensed way, splitting the whole corpus up into four subsets. The figures and percentages given here indicate the average of all files subsumed under one subset.

Table 9: **OV: Subject type: NP vs. pronominal subj.** (Aleksandra corpus: **F18-37**)

***ALEKS* files:** **Age:**	**F. 18-23** **(2;2-2;4)**	**F. 24-29** **(2;5-2;7)**	**F. 30-34** **(2;8-2;11)**	**F. 35-37** **(3;0-3;3)**	**Total** **(2;2-3;3)**
N / NP subject	162 54.9%	126 44.5%	101 40.2%	66 42,30%	455 59,9% av.
Subj. pronoun	94 31.9%	115 40.6%	105 41.8%	83 53.2%	397 59,9% av.
demonstr. pron.	39 13,20%	42 14.8%	45 17.9%	7 4.5%	133 40,1% av.
pronom. subj. total	133 45.1%	157 55.5%	150 59.8%	90 57.7%	530 40,1% av.
Overt subjects total in % (out of fin. utt.)	295 (=44.2%)	283 (=40.6%)	251 (=43.9%)	156 (=35.9%)	985 (=41.5%)

Let's visualize these figures from two perspectives again: the frequency in absolute numbers and in per cent. The two figures below contrast two types of (overt) subjects: NPs (including nouns and proper names) vs. pronominal subjects (including subject pronouns and demonstrative pronouns). Figure 5a below illustrates this contrast for the relevant files (from *Aleks* 2;2-3;1) in absolute numbers, whereas Figure 5b shows the same distribution in per cent:

Figure 5a: Subject Type: Distribution of NPs vs..Pronominal Subjects (*Aleks* 2;2-3;1)

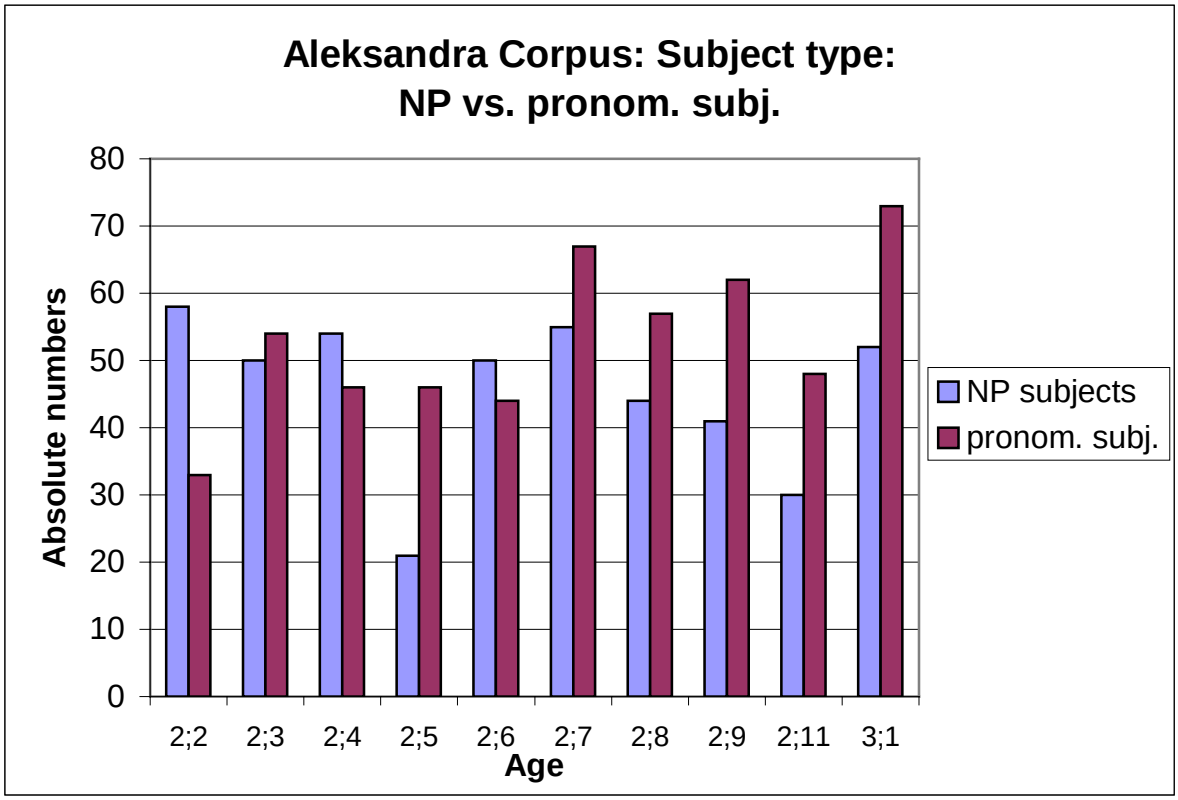
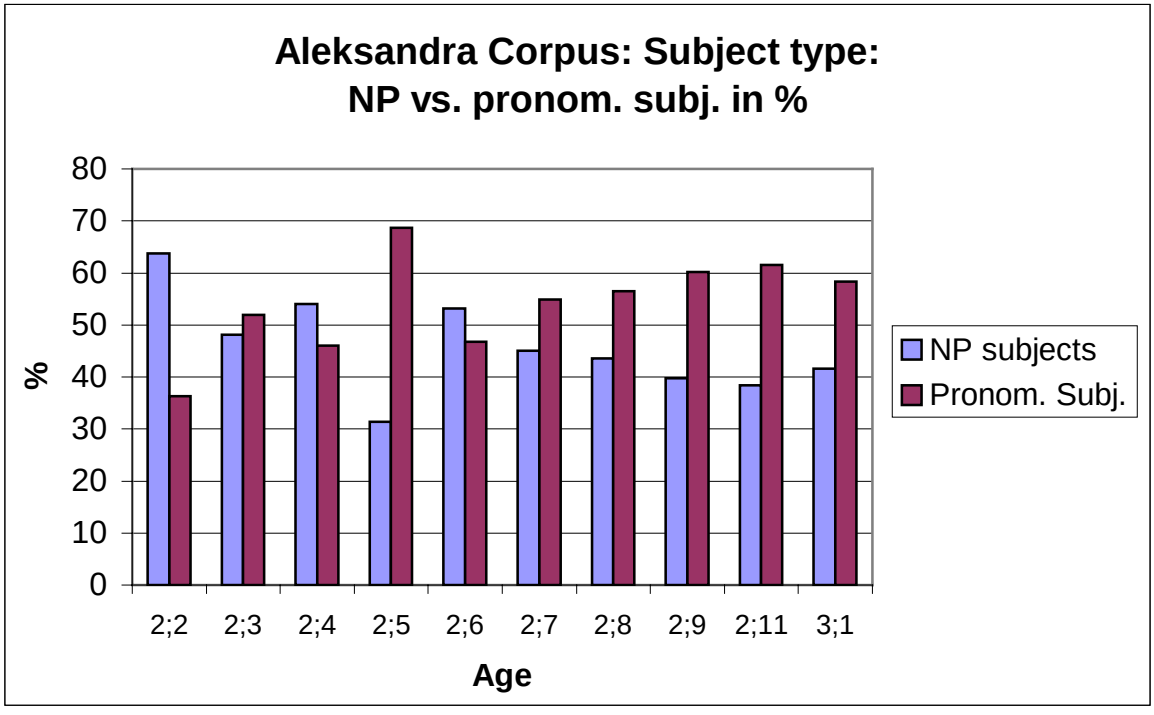


Figure 5b: Subject Type: Distribution of NPs vs. Pronominal Subjects in % (*Aleks* 2;2-3;1)



The tables and figures above reveal some unexpected results with respect to the realization of subjects in Polish child language. First of all, the frequency rates of overt subjects are surprisingly high. Although Aleksandra is acquiring a null-subject language, she uses overt and null subjects at the same time, and both forms are almost evenly distributed. Overt subjects are – almost without exception – used without emphasis or stress, but with neutral intonation, which makes them look like an equal alternative to null subjects in the children's grammatical system rather than a marked form.

From the earliest recordings on, overt subjects may appear both preverbally (16) and postverbally (17):

(16) Tutaj tata śpi.
Here daddy sleep-**3.SG PRES**
'Daddy sleeps here.' [*Aleks* 11, 22 (1;10)]

(17) Tam mieszka babcia.
There live-**3.SG PRES** grandmother
'Grandmother lives there.' [*Aleks* 11, 24 (1;10)]

(16) and (17) come from one of the first quantifiable files (*Aleks* 11, 1;10) and testify to Aleksandra's tendency to experiment with word order variation, a tendency that can be observed throughout the corpus (we will see a few more examples later on).

Let's take a look at the *type* of subject used in early child data. Any utterance with an overt subject was classified according to variable no. 9 "Lexical content (overt) subject" that centres around the general opposition *lexical* vs. *function* words. In the classification of subject types, we distinguish lexical subjects (e.g. Nouns / NPs) from functional structures (personal pronouns or demonstrative pronouns). The results of the counts for each file are shown in the evaluation table in the appendix (including some concrete examples of pronouns found in the given file).

Table 8 and 9 and Figure 5a and 5b show that the rate of lexical subjects (bare nouns, NPs) is strikingly low compared to the amount of pronominal subjects. Throughout the files 18-37, pronominal subjects appear in abundance, as table 8 illustrates very clearly. Among pronominal subjects the most frequent pronouns are personal pronouns (subject pronouns).

Both findings are particularly obvious from File 34 (2;5) on, when frequency rates of pronominal subjects come close to 60% on average.

Apart from the abundance of personal/subject pronouns, the data also display a *variety* of different forms with respect to person, number and gender (for the 3rd-person pronouns), as the data in (18)-(22) illustrate. The first subject pronouns to be found in the *Aleksandra*-Corpus are the pronouns *ty* and *ja*, as listed in the evaluation table in the appendix[3]. They appear in File 15 (2;0.15) for the first time. The following data illustrate the variety of personal pronouns found, some of them even in the earlier files (e.g. (18)):

(18) **Ja** mu zrobię podutkę [poduszkę]
I him do-3.SG PRES-(PERF) pillow
'I will prepare the pillow for him.' [*Aleks* 16, 42 (2;0.29)]

(19) Gołas **ty** jesteś.
naked you-2SG be-2SG PRES
'You are naked.' [*Aleks* 20, 171 (2;3.4)]

(20) **On** się nazywa Dżeki.
He refl.pron. call-3.SG PRES Dzeki.
'He is called Dzeki.' [*Aleks* 18, 162 (2;2.3)]

(21) Cześć, **my** jesteśmy lala i Balbinka.
Hello, we be-1.PL PRES doll and Balbinka
'Hello, we are the doll and Balbinka.' [*Aleks* 20, 85 (2;3.4)]

(22) **One** są nie dobre.
They-FEM be-3.PL PRES not good (=bad)
'They are bad.' [*Aleks* 19, 137 (2;2.20)]

6.3.3 Comparison with Adult Speech (Control-Group)

One question that I have hinted at earlier should be raised now. If we want to attempt an interpretation of the striking frequency rate of pronominal subjects in the Polish child data, we should first check what the distribution looks like in the target or adult system. A comparison between both system seems to be an absolute prerequisite for an evaluation of the above

[3] The very first pronouns in the *Aleksandra*-data are object pronouns: the demonstrative *to* (file 10 (1;09) and strong object pronouns like *nas* 'us' and *ich* 'them' in file 12 (1;10), later followed by weak pronouns.

results. I decided, therefore, to analyse the mother's or adult's speech on a random basis (as a control group) in order to compare the frequency rates found in the child data to the distribution in adult speech. The mother's speech in three representative files was analysed with respect to the distribution of null- vs. overt subjects, and also the type of subject involved in the mother's utterances. Since both the child and the adult data analysed stem from the same recording, thereby sharing the same contextual and situational frame, we have a fairly reliable basis for comparison. The results for the analysis of the adult's speech in files 20, 23, and 35 is given in tables 10 and 11 below.

Table 10: Analysis of Polish **Adult Speech** (Control Group: **Mother):**
Distribution of Null- vs. Overt Subjects

File	Null Subj. Utt.	Overt Subj. Utt.	Fin. Utt. Total
20	164	107	271
	60.5%	39.5%	100%
23	145	102	247
	58.7%	41.3%	100%
35	161	104	265
	60.8%	39.2%	100%
Total :	**470**	**313**	**783**
	60.0% av.	**40.0% av.**	**100%**

The analysis of Polish adult speech in the three files, that were chosen at random, reveals that the mother's utterances show a very consistent distribution of null- and overt subjects. Out of the total of 783 finite utterances analysed from the mother's speech, we find approximately 40% utterances with overt subjects, and 60% null subjects. This fairly stable percentage of overt subjects is still significantly below the corresponding frequency rate for the child in File 20 and especially File 23, when Aleksandra's use of overt subjects is at its peak (53.9%). Towards the end of the recording period, from File 35 on, however, Aleksandra's use of overt subjects decreases considerably, thereby approaching the percentage found in the adult system.

Still, the frequency rate for overt subject utterances in the adult's speech (i.e. 40%) are higher than one might expect from a language that is traditionally classified as a null-subject language. One possible explanation for this fact was pointed out to me by G. Dogil:
The distribution of null- vs. overt subjects in the mother's speech could reveal a current tendency in the adult grammar of contemporary Polish. As G. Dogil (p.c.) points out, present-day speakers of Polish tend to use null- and overt subjects more and more interchangeably. Contrary to accounts in traditional Polish grammars, the grammatical system of present-day Polish seems to be changing from a language where null-subjects are the unmarked case to a system that allows null- and overt subjects to appear in free alternation. If the use of pronominal subjects is no longer associated with emphasis or contrastive focus, the two ways of realizing the subject would be a case of true optionality.
This issue will be picked up in the conclusion that contains some speculation on the diachronic implications of this finding.

The contrast between child and adult speech, however, is much stronger in the distribution of different subject *types*, i.e. N /NP-subjects vs. pronominal subjects, as illustrated in Table 11 below:

Table 11: Analysis of Polish **Adult Speech** (Control Group: **Mother):**
Distribution of Subject Type: N/NP vs. Pronominal Subject

File	**N / NP-Subj.**	**Pronom. Subj.**	**Overt Subj. Total**
20	69	38	107
	64.5%	35.5%	100%
23	58	44	102
	56.9%	43.1%	100%
35	61	43	104
	58.7%	41.3%	100%
Total:	**188**	**125**	**313**
	60.1% av.	**39.9% av.**	**100%**

If we compare these percentages to the child data during the period with the highest frequency of pronominal subjects (i.e. approximately 60%, in Files 24-34), we can conclude that adults almost show the reverse picture: a relation of 40% to 60% (null- vs. overt subjects) in contrast to 60% to 40%, as displayed in Polish child data.

The contrast between adult and child speech decreases, however, as soon as the children start using less overt subjects from File 35 (age 3;0.19) on (in File 35, for instance, only 30%).

The period between the age of 2;5 and 2;11, however, is characterized by an abundance of pronominal subjects in the *Aleksandra*-Corpus that calls for an explanation. We will attempt one by looking for dependencies that can be deduced from results of the data evaluation scheme, - due to the use of interrelated variables, as discussed in section 5.3.3.

6.3.4 Deriving Syntactic Dependencies in the *Aleksandra*-Corpus

Since the percentage of overt subjects, and particularly overt *pronominal* subjects in the *Aleksandra*-Corpus exceeds the corresponding frequency of overt subjects in the adult system, it will be interesting to investigate what factors might cause the excessive use of pronominal subject utterances in Polish child grammar between the age of 2;3 and 3;0. If we attempt to derive syntactic dependencies from the data, the approach using interrelated variables seems promising: it might allow us to establish a correlation between certain grammatical properties of the predicate and the use of null- or overt subjects (see the discussion in section 5.3.3).

First of all, what kind of correlations can be observed between overt subjects and different predicate features?

The first observation is a finding that has been confirmed in many other L1 studies. It concerns the correlation between overt subject utterances and Variable 1 (“Finiteness”) in the data evaluation scheme. If we look at the distribution for this first variable, we can see that overt subjects only occur with finite verbs. Throughout the database, including all the early files, overt subjects *never* appear with non-finite verbs (nor imperatives or bare participles) throughout the *Aleksandra*-Corpus. As a matter of fact, main clause infinitives only occur

marginally in our database, due to the absence of an *Optional Infinitive Stage* in Polish. In the early recordings, however, where main clause infinitives appear a little more frequently (mainly as elliptical structures or associated with volitional meaning), they are never used with an overt subject (without exception, as one can see in the evaluation tables in the appendix).

Our finding is consistent with the observation made in many studies on the *Optional Infinitive Stage* that there seems to be a strong correlation between null subjects and *optional infinitives*.

Another interesting correlation concerns Variable 2 "Verb Type" that classifies finite verbs with respect to the distinction *lexical* vs. *functional* verbs.

The first impressionistic observation that comes to mind when looking at the distribution of null- and overt subject clauses in relation to different verb types in the evaluation tables, is a striking correlation of the *copula* (forms of być 'to be) with *overt* subjects. To find out if this impression can be confirmed by a quantitative data analysis, I re-analysed all instances of the copula in 20 files (*Aleks* 15-34) that cover an age range from 2;0 until 3;0 roughly.

In addition to the counts already shown in the tables for each file individually, I grouped five files together at a time to show the overall development. Table 12 below summarizes the results in a condensed form, showing the distribution in absolute numbers and column percentages:

Table 12: **Correlation Between Copular Utterances and [+/- Overt Subj.] Utt.**
Aleksandra-Corpus: File 15-34 (2;0 – 3;0)

***Aleks*- files:**	**F. 15-19**	**F. 20-24**	**F. 25-29**	**F. 30-34**	**Total**
Age:	**(2;0-2;2)**	**(2;3-2;5)**	**(2;5-2;7)**	**(2;8-3;0)**	**(2;0-3;0)**
Cop. / Null-subj.	27	15	20	28	90
	36.9%	14.6%	20%	19.4%	21.4% av.
Cop./ overt subj.	46	88	80	116	330
	63.1%	85.4%	80%	80.6%	78.6% av.
Total:	73	103	100	144	420

Table 12 shows indeed a striking correlation between copular utterances and an overt realization of the subject. With frequency rates going up to 85% in the second data subset (2;3-2;5), and an average rate of 78.6% (2;0-3;0), this finding seems to represent a solid child language phenomenon that is limited to this developmental period. Later on, however, starting with file 35 at the age of 3;0.19, we find an even distribution of copular utterances with null- and overt subjects respectively: in File 35, the distribution is 9 null-subj.- vs. 5 overt-subj. copular utterances, in File 36, the relation is 12:12, and in the last file 37, it is 23:23 (see data evaluation tables in the appendix). That shows that from the age of 3;0 on, the child's use of copular constructions approaches the adult system.

As for further differentiation regarding the *type* of subject involved, I found a strong preference for the use of copular utterances with *pronominal* subjects, as illustrated in the following data from the *Aleksandra*-Corpus. Pronominal subjects, including personal subject pronouns and demonstrative pronouns, occur in the total set of copular clauses at a frequency of 81.7% (as opposed to 18.3% lexical subjects, i.e. Nouns / NPs). The data (23) – (27) show personal pronouns, whereas (27) illustrates the frequent use of the demonstrative pronoun *to:*

(23) **Ja** jestem malutka.
I be-1 SG PRES little.
'I am little.' [*Aleks* 20, 84 (2;3)]

(24) **Ty** jesteś Tólewna Śnieżta [Królewna Śnieżka]
you-2 SG be-2 SG PRES Snow Queen
'You are the Snow Queen.' [*Aleks* 23, 115 (2;4)]

(25) **Ona** jest duża.
She-3 SG-FEM be-3 SG PRES big
'She is big.' [*Aleks* 23, 123 (2;4)]

(26) **One** są nie dobre.
They-FEM be-3.PL PRES not good (=bad)
'They are bad.' [*Aleks* 19, 137 (2;2.20)]

(27) **To** jest sweterek.
That be-3 SG PRES jumper-DIM
'That is a little jumper.' [*Aleks* 23, 44 (2;4)]

This striking correlation between copular utterances and overt subjects on the one hand, and – with respect to subject type – pronominal subjects on the other hand is supported by similar findings for child data from a few other languages. In the next section, I will present some

interesting parallels that show up in the analysis of Croatian, Italian and Austrian German child data.

Before turning to the cross-linguistic level, however, let's look for some further interrelation between the excessive use of pronominal subjects and other – possibly interacting – developmental tendencies during the same period. Again, we are well advised to study the results of the data evaluation scheme to profit from our longitudinal database. As discussed in chapter 5, only longitudinal studies allow the researcher to compare critical data with material from preceding or subsequent stages, as well as with simultaneous trends in other areas of language development.

So, - which other (simultaneously occurring) factors could the use of pronominal pronouns be attributed to?

The copular results already point to one direction: copular constructions and functional elements (such as subject pronouns) seem to cluster together.
Throughout the files studies here, all child data display correct agreement between the grammatical subject and the copular verb. This is in so far worth mentioning as the copular verb *be* does not assign a thematic role (see Moro (1995, 1997)); consequently the child cannot rely on a semantically based approach to establish agreement between the copula and the structural subject. Instead, she can only rely on structural knowledge that is manifested in early mastery of morphosyntactic features of the target language.

When investigating which other acquisitional processes can be observed during the period of preferred use of pronominal subjects, we can spot a few new trends in the development of tense and verb morphology. At the same time when pronominal subjects begin to appear more and more frequently (i.e. from file 19/20 (2;2) on), we can also observe the appearance of new tense forms (past and future tense) and more variety in the present tense paradigm (increasing use of plural inflections). In general, this period seems to be characterized by *overt* morphological expression of *functional* features, as opposed to the earliest stage of acquisition (showing the phenomenon of *Descriptive Imperatives,* see section 6.2), where morphosyntactic features were left unexpressed or neutralized, resulting in the use of *surrogate verb forms* (6.2.4). As for the excessive use of pronominal subject (without emphasis or any other related function), the child seems to have gone to the opposite end of

the continuum: an "over-expression" of (partly redundant) morphological features that are not required by the target language, since the Polish system of verbal inflection shows sufficient distinctions to identify person and number even in the absence of subject pronouns. Crucially, I do not mean to say that the *underlying* grammatical system of the child has changed. We have seen independent evidence for morphosyntactic knowledge and the presence of functional categories in early child grammar. What has changed, however, is the way of *overtly* expressing this knowledge, probably due to a more efficient process of *accessing* morphological knowledge (along the lines of Phillips' (1995) approach).

I would like to add a brief comment on the development of some deviant word order patterns, that might shed additional light on the acquisitional mechanisms that show up in the *Aleksandra*-Corpus.

As for word order patterns, we can notice a strong tendency to experiment with word order variation that is manifested in a variety of word order patterns. These word order "experiments" found in the *Aleksandra*-data often go beyond the limits of permissible syntactic variation, even in a language like Polish that enjoys reasonable freedom in word order.[4] Data (28), (29) and (30) illustrate Aleksandra's early word order experiments and her use of discontinuous constituents (printed in bold) which are hardly acceptable in the target language (G. Dogil, p.c.). These deviant structures are quite frequent: in each file, there are at least two or three such examples, displaying a rich variety of word order patters. The following data occur side by side with utterances that show standard word order patterns where the elements printed in bold would be realized adjacent to each other:

(28) **Muszę** ja **trzymać** tą jaję [lalę] [declarative context!]
must-1 SG I hold-INF this doll-AKK
'I must hold this doll.' [*Aleks* 20, 83 (2;3)]

(29) A **ten** idzie **wąż** z nogami, (tup tup tup…)
And this go-3 SG snake with legs
'And this snake goes with legs.' [*Aleks* 20, 58 (2;3)]

(30) To jest **materac** z myszką **mój.**
this be-3 SG PRES mattress with mouse my
'This is my mattress with the mouse.' [*Aleks* 34, 113 (2;10)]

[4] I am grateful to G. Dogil for discussing many critical data with me, as well as for giving me his grammaticality and acceptability judgements for the deviant word order patterns.

(31) Tu **będzie** - wiesz – tata **mieszkał** w biurze.
Here aux-3SG FUT -you know- daddy live- PAST PARTIC in office
'You know, daddy will live here in the office.'
[*Aleks* 16, 154 (2;0)]

The use of discontinuous constituents in Aleksandra's early stages of language acquisition is particularly interesting, as these patterns provide ample evidence for differentiated knowledge of structural relations, especially in cases like (28) - (31), where the child cannot rely on any linear relation between two adjacent elements. Therefore, these data can only be explained in terms of syntactic knowledge. The fact that structures like (28)-(30) are ungrammatical in the target language or at least stylistically highly marked, makes this type of data even more intriguing: they allow us to gain an insight into how the child actually proceeds: on her way to the acquisition of a grammatical property, the child seems to be exploring various structural options provided by UG which (temporarily) leads to coexisting alternative implementations of this grammatical property in the child's interim grammar. There seem to be grammatical properties of the target language that are not acquired strictly incrementally but competitively, i.e. by eliminating these coexisting alternative implementations (see also Schaner-Wolles (1995/96) for an analysis of the acquisition of negation in this framework).

Impressionistically speaking, it looks like the child's acquisitional mechanisms that are revealed in data like (28)-(31), match an L1 approach along the lines of Lebeaux (1988), Verrips (1994), LeBlanc (1995), van Kampen & Evers (1995) and van Kampen (1997). These authors have tried to extend the parametric model in order to explain optionality and variation. They have proposed an alternative scenario of the acquisition process by suggesting that the acquisition of a given property of the target language involves successive elimination of universally available options, as summarized in Schaner-Wolles (1995:1):

> The acquisition of a property P of the grammar of a language L consists in the implicit elimination / suppression of universally available implementation options for P which are not instantiated in the grammar of L.

Under this view, parameter setting does not equal the setting of cognitive "switches", but rather equals the elimination / suppression of alternative options.

6.3.5 The Realization of Subjects: A Cross-Linguistic Comparison

Let's now to the cross-linguistic level of analysis that is particularly relevant to the discussion of null subjects in early child grammar. As I already mentioned at the beginning of this chapter, this topic is currently under much debate. One of the most salient properties of early grammatical systems is the apparent optionality of subjects. Acquirers of all languages (even of target languages that do not permit null subjects) seem to pass through a stage in which they frequently omit subjects. This is illustrated in the following subjectless utterances from English, Dutch and French, all languages which do not license null subjects in their adult form:

(32) a. Want to get it.
b. Not making muffins.

(33) Ook toren bouwen. [Dutch]
also tower build
'(I) build a tower too.' [Haegeman (1994)]

(34) Veux pas lolo. [French]
want not water
' (I) don't want water.' [Pierce (1992)]

Based on data like (32)-(34), many studies (following Hyams (1986)) have suggested that children acquiring non-null-subject languages have mistakenly set the Null Subject-/pro drop-Parameter to the Italian value. At first sight, this hypothesis seems to be appealing since it relates the stages of acquisition directly to the typological variation found in adult languages, and thus explains the developmental stages directly in terms of independently motivated principles and parameters of grammar.

On the other hand, the *Parameter Mis-Setting Hypothesis* of early null subjects (Hyams (1986), has been challenged on both conceptual and empirical grounds. One major conceptual problem consists in the difficulty to explain the re-setting of the parameter that was mis-set initially given the fact that the child's input never changes. From an empirical point of view, the hypothesis was challenged on the basis of structural and distributional differences between early null subjects and null-subjects in adult pro-drop languages. It was found out that early

null subjects are impossible or highly infrequent in the following contexts: questions with a fronted wh-element (see Valian (1990)), subordinate clauses (Valian (1990), Weissenborn (1992)) and main clauses with fronted XPs other than the subject. Italian children, in contrast, seem to use null subjects essentially the way Italian adults do, i.e. in simple finite clauses (35), but also in finite subordinate clause (36), as illustrated in the following data (Guasti (1993/4)):

(35) E mia gonna.
Is-3 SG my skirt
'(It) is my skirt.'

(36) Mama dice che non è simpatico.
Mama say3 SG that not is nice.
'Mama says that (it) is not nice.'

In early English, in contrast, null subjects are restricted to main clauses. If English-speaking children initially have the same grammatical system as a null-subject language, we would expect them to use null subjects with the same frequency and distribution as Italian adults. This is not the case, however, as the distributional facts mentioned above demonstrate. Moreover, Valian (1991) observes that while Italian children omit subjects at a rate of about 70% of the time (about the same rate as Italian adults), English-speaking children drop null subjects at a rate of 30-50%.[5]

This takes us back to the question of cross-linguistic evidence for the realization of subjects. An interesting language to look at is Russian since, as Bar-Shalom & Snyder (1998) put it, Russian and Polish are morphologically similar, and yet form a "minimal pair" with respect to their pro-drop status: Polish, but not Russian, is a pro-drop language.

Let's look at the percentages of null- vs. overt subjects in Russian child language then. In their recent publication, Bar-Shalom & Snyder (2001, table 2) give the developmental pattern for overt subjects in the *Svetlana*-Corpus, which ranges from 11% (1;8-1;11) to 29% (2;4-2;5). In the last file, Svetlana used overt subjects at a rate of 34%. Table 1 in the same study compares the distribution of null- vs. overt subjects in the speech of Svetlana at the age of 2;5

[5] Another theoretical problem of Hyams' original proposal was how the content of early null subjects is recovered. While null subjects are identified by Φ-features expressed on the verb in pro-drop languages, such a mechanism is not available in early non-pro-drop languages like English, due to the lack of a rich agreement system. Consequently, Hyams (1992) retracted her pro-drop proposal and suggested instead a topic-drop account, claiming that these early null subjects are of the Chinese kind and do not depend on Φ-features for their identification. Without going into details here, this account is also problematic w.r.t. the distributional facts.

to the percentages in her mother's speech during the same recording. The figures are reproduced below (table 13, Bar-Shalom & Snyder (2001:97):

Table 13: Distribution of **overt vs. null subjects** in the speech of Svetlana and her mother (when Svetlana was 2;5)

	Overt Subject	**Null Subject**
Svetlana:	29.3% (12 / 41)	70.7% (29 / 41)
Mother:	47.3% (18 / 38)	52.6% (20 / 38)

As shown in Table 13, Svetlana's mother produced overt subjects more often than did Svetlana at age 2;5, and yet, the percentage of overt subjects seems to be relatively low if we consider that Russian is a non-pro-drop language. However, the low rates could be due to the fact that Russian still allows limited, discourse-related subject drop (Bar-Shalom & Snyder (2001), footnote 6). Unfortunately, the total number of utterances analysed is very low, so that one cannot really draw any general conclusions.

As for the high percentage of null subjects in the child data, this is in fact not a surprising finding given the fact that Russian does show an *Optional Infinitive Stage,* and this stage tends to always correlate with null subjects.

The distribution for Russian can be related to similar findings in various other languages, such as Dutch, Flemish, French and German, that are summarized in the following table (see next page) for a better comparative overview:

[Sources of the figures in table 14: Flemish and French data from Krämer (1994), Dutch data from Haegeman (1994), German data from Behrens (1993).]

Table 14: Distribution **Overt vs. Null-Subjects in Finite and Non-finite clauses in %**

	Finite Verb		Non-Finite Verb	
%	**Overt subj.**	**Null subj.**	**Overt subj.**	**Null subj.**
Dutch	**68**	**32**	**15**	**85**
Flemish	**75**	**25**	**11**	**89**
French	**74**	**26**	**7**	**93**
German	**80**	**20**	**15**	**85**

Table 14 shows that in overt subject languages such as Dutch, Flemish, French and German, the realization of the subject as overt or null depends on the finiteness of the clause. Thus, main clause infinitives tend to occur with null subjects, while finite sentences typically contain overt subjects. The correlation is not perfect, however, as the figures show: null subjects do occur with finite clauses, and overt subjects with main clause infinitives, but the percentages are far less than would be expected by chance.

Not all data from overt subject languages, however, follow the same distributional pattern, as longitudinal data from Austrian German (the *Nico*-Corpus[6]) show.
In Katičić and Schaner-Wolles (2001) the *Nico*-data (files between 2;2 and 2;9) are analysed with respect to the distribution of [+/- overt subject] and the variable subject type, both for finite and for non-finite verbs. In finite clauses, Nico (between 2;2 and 2;9) produces overt subjects at a frequency rate of 77% on average (and, consequently, 23% for null subjects). In non-finite main clauses, on the other hand, the rate of overt subjects is at 44% between 2;2 and 2;5, and 40% between 2;5 and 2;9, whereas null-subjects are used in the two stages at 56% and 60% respectively. The distributional contrast between finite and non-finite verbs with respect to the realization of subjects is not as sharp in the *Nico*-Corpus as has been claimed for other non-null-subject languages (cf. table 14 above). Yet, Nico produces far more null subjects in main clause infinitives than in finite clauses.

[6] The Nico-data were evaluated in the cross-linguistic L1-project "Erstsprachlicher Grammatikerwerb des österreichischen Deutsch im Vergleich", run by Chris-Schaner-Wolles at the University of Vienna. For a comprehensive analysis of the Nico-data with respect to finiteness and verb position, see Köhler (1998).

As far as the distribution for subject *type* is concernced, Nico is reported to use pronominal subjects in finite clauses at a rate of 26-29% between 2;2 and 2;9.
Another interesting recent finding concerns Nico's distribution of subject types in copular sentences: he uses pronominal subjects at a frequency of 83.9%, while only 16.1% of Nico's copular constructions contain a lexical subject (N/NP) (see Czinglar et al. (2003)).

This result is very similar to our finding in the *Aleksandra*-data (see section 6.3.4), where pronominal subjects, including personal subject pronouns and demonstrative pronouns, occur in the total set of copular clauses at a frequency of 81.7% (as opposed to 18.3% lexical subjects, i.e. Nouns / NPs).

This striking frequency in the use of pronominal subjects in early child language has been confirmed by two further studies based on Croatian and Italian child data[7].
Let's first look at the Croatian data, the *Antonija*-Corpus (cf. section 6.2.4), which is also analysed in Katicić and Schaner-Wolles (2001).
In finite clauses, the percentage of overt subjects begins at 18% at the age of 1;7 as opposed to 82% null subjects. In the following files, the rates for overt subjects go up to 30% at 2;1 (with 68% null subjects). For the Antonija-files between the age of 1;7 and 2;1, Katicić (p.c.) calculates an average percentage of 34% of overt subjects, as opposed to 66% null subjects.

As for the *type* of subject involved, Antonija shows a parallel preference for pronominal subjects – compared to the Polish child Aleksandra. Between 1;7 and 2;1, 70% of Antonija's overt subjects utterances contain a subject pronoun, while only 30% have a lexical subject (N/NP) (Katicić p.c.).

Finally, we will look at some Italian data that are analysed in the framework of the L1-project at the University of Vienna (see footnote 7). These data analysed by Livia Tonelli complement our cross-linguistic survey in a nice way, since they do not only give the distribution of overt vs. null subjects, but analyse the set of overt subjects with respect to the *type* of subject as well. The study is based on the *Marco*-Corpus that consists of 27 files (recorded at two-weeks-intervals) between the age of 1;5 and 2;5, containing a total of 1065

[7] Part of these studies is still work in progess in the framework of the cross-linguistic L1-project "Erstsprachlicher Grammatikerwerb des österreichischen Deutsch im Vergleich", run by Chris-Schaner-Wolles, where I was involved as one of the external cooperation partners for the comparative section that deals with a number of languages, among them Croatian (A. Katicić), Italian (L. Tonelli) and Polish.

utterances. 78% of Marco's finite sentences are null-subject utterances, while 22% contain an overt subject. Between 1;5 and 2;5, Marco produced a total of 234 overt subjects, out of which 170 are pronominal (which equals 72.6%) and only 64 are nominal subjects (41 NPs and 23 proper nouns), which corresponds to 27.4% lexical subjects (Chris Schaner-Wolles, p.c.). Thus, the Italian *Marco*-data match the Polish and Croatian data very well, lending additional support to the observation that young children acquiring a null-subject language show a strikingly high frequency rate for overt pronominal subjects.

6.4 Conclusion

In this chapter, I have presented the major results of the quantitative and qualitative analyses of my database, according to the three principal research questions (Q1-3) formulated at the end of chapter 5. The first research topic (Q1) concerns the distribution of finiteness in all three corpora, which is discussed in section 6.1. In the cross-linguistic framework of current language acquisition studies, the aim of this study is to address the question if there is any evidence for an *Optional Infinitive Stage* in Polish. Both the quantitative analysis presented in section 6.1.2, and the qualitative analysis – aiming at the interpretation of main clause infinitives – (6.1.3) point to a negative answer: main clause infinitives mainly occur (if at all) in modal, volitional or elliptical contexts, while infinitval *forms* do occur much more frequently inside [aux./modal + infinitive]-constructions, thereby rendering a phonological explanation for their low frequency very unlikely.

Instead, an investigation of the earliest stages of acquisition, the one- and two-word stage (section 6.2, cf. Q2) has revealed an interesting, unexpected finding: the imperative (and *not* the infinitive) seems to play a special role in the children's early grammatical system. Until the age of 1;8 approximately, the imperative is used in declarative contexts, and is mostly associated with a descriptive meaning ("*Descriptive Imperative"*). At the same time, the Polish children analysed here seem to have a very good command of subject-verb agreement and other items of morphosyntactic knowledge, as shown in section 6.2.3. Having considered some cross-linguistic evidence for a similar kind of "default" use of *Descriptive Imperatives* in Russian child language (6.2.4), I suggest that this deviant use of the imperative form is not

due to lack of morphosyntactic knowledge (for which, again, I have sufficient independent evidence), but rather to lack of an automated process of *accessing* that knowledge. In case of failure to retrieve the relevant morphological realization, the child uses the imperative as a *surrogate verb form* instead (at the earliest stage of acquisition).

The third research topic (spelled out in Q3) focuses on the realization of null subjects in the *Aleksandra*-Corpus and is also related to the issue of *optionality* in early child grammar. One of the most salient properties of early language is the apparent optionality of subjects. Children acquiring a non-null-subject language have been claimed to pass through a stage in which they frequently omit subjects. In section 6.3.1, I address this issue from the other side, trying to establish the distribution of null- vs. overt subjects in a language that is traditionally classified as a null-subject/pro-drop language. In the *Aleksandra*-data, I find a high frequency of overt subjects (well above 40% on average), particularly between the age of 2;2 and 2;11, when the percentage of overt subjects goes up to 53.9% at one point.

Given this surprisingly high frequency of child utterances with overt subjects, it seems worthwhile to investigate the *type* of subject involved (lexical, i.e. N/NP, vs. pronominal (personal or demonstrative pronouns)). The distributional analysis reveals that the rate of nominal/lexical subjects is strikingly low compared to the abundance of pronominal subjects (notably personal pronouns). In order to check if this phenomenon is restricted to child language or if it appears in adult Polish as well, I analysed the distribution of overt vs. null subjects in the mother's speech on a random basis in section 6.3.3. A comparison between both systems shows that the adult's percentage of overt subjects is significantly below the corresponding frequency rate for the child at the crucial age (2;4), while Aleksandra's use of overt subjects decreases considerably in the later files and approaches the frequency rate found in the adult system.

Section 6.3.4 makes an attempt at deriving syntactic dependencies in the *Aleksandra*-Corpus in order to find out what factors might cause the excessive use of pronominal subject utterances in Polish child grammar between the age of 2;3 and 3;0. The most important correlation we found was between copular utterances and an overt realization of the subject, which shows that copular constructions and other functional elements (here: subject pronouns) seem to cluster together. Based on observations from other areas of language development during the given period (e.g. (partly ungrammatical) word order experiments,

use of discontinuous constituents etc.), we have seen evidence for coexisting alternatives in the child's grammatical system. This could be interpreted along the lines of L1 approaches that have proposed a scenario of the acquisition process that involves successive elimination of universally available options. According to this view, there are grammatical properties of the target language that are not acquired strictly incrementally, but competitively, i.e. by eliminating coexisting alternative implementations of the grammatical property.

Finally, section 6.3.5 turns to the cross-linguistic level of the analysis, discussing data from different non-null-subject languages (Dutch, Flemish, French, German and Russian) as well as from null-subject languages, such as Croatian and Italian. After discussing the *Parameter Mis-Setting Hypothesis,* that was put forward in order to explain early (ungrammatical) null subjects in overt subject languages, I compared the results from all these studies. It turns out that in overt subject languages, the realization of the subject as overt or null depends on the finiteness of the clause: main clause infinitives tend to occur with null subjects, while finite sentences typically contain overt subjects.

As for null subject languages, the frequencies of overt null subjects in Italian and Croatian child language are lower than in Polish, but all three languages show a striking parallel: a percentage of pronominal subjects, in other words: a strong preference of *pronominal* over *nominal* subjects. This phenomenon is surprising since pronouns are usually reported to be acquired late (cf. Haag-Merz (1995) for Swabian). The fact that the children prefer a functional to a lexical element of grammar coincides with a period of explorations into the options of their target language (as shown by the word order "experiments" in (28)-(31)).
All phenomena analysed so far speak in favour of differentiated morphosyntactic knowledge and the presence of functional categories in the early child grammar.

Chapter 7: Polish Child Data in a Cross-Linguistic Context

Which conclusions can be drawn from the analyses presented in chapter 6?
In this final chapter, I will summarize the main results of the quantitative and qualitative analyses and discuss them from a cross-linguistic perspective. I will begin by reviewing the three main research topics of this thesis and recapitulate the results for the individual research questions. Besides, I will discuss the results with respect to their implications for the theoretical questions to be addressed in this study.

The starting point of the analysis was a well-known cross-linguistic observation in the context of the *Optional Infinitive* debate: the phenomenon of *optionality* in the grammatical system of young children. Based on a number of comparative acquisition studies that are reviewed in chapter 3, we raised the question of children's clausal architecture by considering two conflicting hypotheses: the *Small Clause Hypothesis* (defending the view that early clauses are mere instantiations of VP, thus lacking functional categories, in particular the inflectional category IP), and the *Full Competence Hypothesis* that assumes adult-like structural representations in early grammars, thereby positing the presence of functional categories.

Empirical evidence for the latter comes from first language acquisition studies that apply the method of comparative analysis, as it is used for the study of word order phenomena in adult systems, to child data. If we analyse the distribution of verbs with respect to negation in French, for instance, we can see a striking correlation between finiteness and verb position with respect to negation: children correctly place a finite verb before the negation and an infinitive after the negation. These findings (and similar distributional results for German with respect to Verb-Second position) seem to point to the presence of functional categories and the validity of the *Full Competence Hypothesis.*

This kind of evidence speaks in favour of a *Continuity View* of language development which amounts to claiming that children's grammars include the same functional projections as adults' grammars, and that the early clausal architecture of the child's grammatical system corresponds to the adult one.

This view is challenged by the phenomenon of *optionality* in early child grammar. The phenomenon under consideration – the alternation of finite verb forms with an *optional* (main clause) *infinitive* at an early age of approximately 2;0 – can be found in many – but not all – languages. Since these main clause infinitives ("*optional infinitives"* in Wexler's approach) are ungrammatical in the target language, they represent a case of *discontinuity*, thereby challenging the *Continuity View* scenario.

As Hyams (see full version of the quotation at the beginning of chapter 1) points out, the phenomenon of *optionality* also raises theoretical problems with respect to linguistic theory:

> Despite the apparent pervasiveness of optional rules in early grammar, optionality is rather mysterious. Why is a rule that is obligatory in the adult grammar optional for the child? Optional rules raise problems both from a linguistic-theoretic perspective and from a learnability perspective.
> Linguistic theory has moved away from the optional rules of the Standard Theory (Chomsky 1965). Within current theory, optional processes are ruled out by licensing principles or by economy considerations (Chomsky 1992).

Both theoretical and acquisitional research has come up with a number of proposals and theories to account for *optional infinitives* in early child grammar. We have looked at *Underspecification Accounts*, the *Truncation Account* and also one account that analyses main clause infinitives as finite structures (Phillips (1995, 1996)).

In my opinion, the phenomenon of *optionality* can be best addressed in its cross-linguistic dimension. That is why this study attempts to contribute to the cross-linguistic discussion by providing empirical evidence from Polish, a rich-agreement language where the infinitive constitutes a highly marked form within the paradigm.

My analysis of the distribution of finiteness (6.1.2) has shown that in all three corpora, there seems to be no evidence for a genuine *Optional Infinitive Stage* in Polish. The data were analysed both quantitatively (section 6.1.2) and qualitatively (6.1.3). As for the latter, we analysed the main clause infinitives in the two corpora where MCIs occurred most frequently – the *Dagmara*- and the *Anna*-Corpus – with respect to their intended meaning. The context analyses for both corpora point to two observations:

On the one hand, the control counts show that infinitival forms do occur in "complete" [auxiliary/modal + infinitive]-structures in both corpora, even in the *Anna*-Corpus that shows

an extremely low percentage of MCIs. This fact provides counterevidence to the idea that infinitives may be absent due to pronunciation difficulties. The data, however, clearly show that infinitives do occur as part of "complete" [modal + infinitive]-constructions and that they even do so with considerable frequency (again, especially in the *Dagmara* data).

Secondly, and more crucially, the data seem to indicate that infinitives in main clauses mainly occur in modal contexts. These root infinitives are often used in utterances that represent answers to a question containing a modal or auxiliary verb. That means that such structures tend to be pragmatically licensed and represent, therefore, possible grammatical structures of the target language as well. The majority of main clause infinitives occurring in modal contexts involve volitional or future tense utterances in which the modal or auxiliary appears to have been dropped. Such elliptical utterances are also possible in the adult system, and, consequently, do not represent cases of *optional infinitives.*

One possible explanation might be hat we simply "missed" an earlier period of use of non-finite main verbs. In order to check this possibility, we extended the database, that originally consisted of the *Dagmara*- and *Anna*-Corpus only, to the *Aleksandra*-corpus which covers the whole range of acquisitional stages, starting from the utterance of single words up to and including multi-word utterances.

The qualitative analysis of the earliest stages of Aleksandra's language development has revealed a surprising result: instead of infinitives, there is a frequent use of imperative forms. Even in contexts where the infinitive is required in the target language (for example in answers to questions containing [modal verb + infinitive]-constructions), Aleksandra still uses imperatives:

(7)	M:	Co Michał będzie robi**ć?** What Michael will do-INF 'What will Michal do?' -	[mother's input]
	A:	Cita**j** read-**IMP** 'Read!'	[*Aleks* 5, 9 (1;06)]

We can safely conclude, therefore, that even at this early stage in the acquisition of Polish, the infinitive does not occur in the way it has been claimed for many other (OI-)languages.

Instead, the imperative seems to play a special role in these earliest stages of acquisition, until the end of the two-word stage roughly, i.e. until the age of 1;8. Crucially, the imperative is used in declarative contexts, and is mostly associated with descriptive meaning We may tentatively suggest that the imperative seems to represent a kind of default form for the acquisition of verbal inflection in Polish in these earliest stages. Later on, i.e. from the age of 1;9 on, this option of imperatives used in declarative contexts (*"Descriptive Imperatives")* seems to be gone.

My findings for Polish have been replicated for Russian in a study by Bar-Shalom & Snyder (2001) of a very young child (Svetlana) who was also shown to use *"Descriptive Imperatives"*.

The Polish and Russian results are supported by similar findings in other languages, notably pro-drop / null-subject languages. In Croatian child language data, for instance, there seems to be evidence for an overgeneralized use of both imperatives and 3rd person-singular forms for various verbal categories in the earliest stages.

Similar findings are reported in Dressler & Makovec-Cerne (1995) for the early stages in the acquisition of Slovenian, when the child initially uses imperative forms in declarative contexts, parallel to my findings for Polish.

Considering all these results from different languages, how could the phenomenon of *Descriptive Infinitives* and other "overgeneralized" forms be captured in a unified account?

Bar-Shalom and Snyder (2001) suggest that both Polish and Russian *Descriptive Imperatives* could result from the children's misanalysis of the imperative as a bare stem. On this view, the Bare Stem Parameter (cf. Hyams (1986)) would sometimes remain unset, at early ages, for Russian and Polish children. Hence, the Russian and Polish children would initially mistake the imperative for a bare stem, and would mistakenly believe that bare stems are morphologically permissible as surface forms, resulting in *Descriptive Imperatives.*

Any account of this phenomenon, however, has to consider the fact that children acquiring different rich-agreement languages (including Italian, Russian and Polish) were shown to be

able to distinguish between finite and non-finite verbs, and, thereby, were shown to possess morphological and syntactic knowledge from the earliest stages of language acquisition on. What might be missing initially, however, is a fast, automated process of *accessing* that knowledge.

By differentiating between very young children's nearly error-free morphological knowledge and the less reliable *implementation* of that knowledge, we can reconcile two apparently conflicting findings and gain a unified account. Therefore, I would take these *Descriptive Imperatives* to be a natural "default" form that is used when the child fails to retrieve the relevant morphological realization. During the earliest stages of language acquisition, the child would use the imperative as a "surrogate" verb form whenever the features inserted in the inflectional system cannot otherwise be expressed. Their semantic interpretation, however, remains unchanged, even when their morphosyntactic features have been neutralized.

Support for this hypothesis also comes from Austrian German (see Schaner-Wolles (2000)), i.e. the Nico-data that show an *Optional Infinitive Stage*. Contrary to what has been claimed in the literature for OI-languages, Nico uses MCIs with temporal reference (present, future, past), even with modal verbs. A considerable percentage of his MCIs, however, have to be interpreted as "surrogate" verb forms, similar to the Polish and Russian *Descriptive Imperatives.*

Further support for this hypothesis comes from languages that have no infinitives in their grammatical systems, such as Greek (use of participles, see Varlokosta, Vainikka & Rohrbacher (1996)), and Inuktitut (widespread use of participial default forms, see Crago & Allen (1994)). While infinitives are the default verbal forms most commonly used in Western European languages, there is no reason why they should be the privileged default forms across all languages.

As for our third research topic, my analysis has revealed some unexpected results with respect to the realization of subjects in Polish child language. First of all, the frequency rates of overt subjects are surprisingly high. Although Aleksandra is acquiring a null-subject language, she uses overt and null subjects at the same time, and both forms are almost evenly distributed. Overt subjects are – almost without exception – used without emphasis or stress, but with

neutral intonation, which makes them look like an equal alternative to null subjects in the children's grammatical system rather than a marked form.

The distribution of null- vs. overt subjects in the mother's speech could reveal a current tendency in the adult grammar of contemporary Polish. As G. Dogil (p.c.) points out, present-day speakers of Polish tend to use null- and overt subjects more and more interchangeably. Contrary to accounts in traditional Polish grammars, the grammatical system of present-day Polish seems to be changing from a language where null-subjects are the unmarked case to a system that allows null- and overt subjects to appear in free alternation. If the use of pronominal subjects is no longer associated with emphasis or contrastive focus, the two ways of realizing the subject would be a case of true optionality.
As Ian Roberts pointed out to me, this result matches the observation that null-subjects can be lost diachronically. – Is Polish on the way to an overt subject language? We will leave this issue open.

The third research topic (spelled out in Q3) focuses on the realization of null subjects in the *Aleksandra*-Corpus and is also related to the issue of *optionality* in early child grammar. One of the most salient properties of early language is the apparent optionality of subjects. Children acquiring a non-null-subject language have been claimed to pass through a stage in which they frequently omit subjects. In section 6.3.1, we addressed this issue from the other side, trying to establish the distribution of null- vs. overt subjects in a language that is traditionally classified as a null-subject/pro-drop language. In the *Aleksandra*-data, we find a high frequency of overt subjects (well above 40% on average), particularly between the age of 2;2 and 2;11, when the percentage of overt subjects goes up to 53.9% at one point.

Given this surprisingly high frequency of child utterances with overt subjects, it seems worthwhile to investigate the *type* of subject involved (lexical, i.e. N/NP, vs. pronominal (personal or demonstrative pronouns)). The distributional analysis reveals that the rate of nominal/lexical subjects is strikingly low compared to the abundance of pronominal subjects (notably personal pronouns). In order to check if this phenomenon is restricted to child language or if it appears in adult Polish as well, we analysed the distribution of overt vs. null subjects in the mother's speech on a random basis in section 6.3.3. A comparison between both systems shows that the adult's percentage of overt subjects is significantly below the

corresponding frequency rate for the child at the crucial age (2;4), while Aleksandra's use of overt subjects decreases considerably in the later files and approaches the frequency rate found in the adult system.

Section 6.3.4 made an attempt at deriving syntactic dependencies in the *Aleksandra*-Corpus in order to find out what factors might cause the excessive use of pronominal subject utterances in Polish child grammar between the age of 2;3 and 3;0. The most important correlation I found was between copular utterances and an overt realization of the subject, which shows that copular constructions and other functional elements (here: subject pronouns) seem to cluster together.

Based on observations from other areas of language development during the given period (e.g. (partly ungrammatical) word order experiments, use of discontinuous constituents etc.), we have seen evidence for coexisting alternatives in the child's grammatical system. This could be interpreted along the lines of L1 approaches that have proposed a scenario of the acquisition process that involves successive elimination of universally available options. According to this view, there are grammatical properties of the target language that are not acquired strictly incrementally, but competitively, i.e. by eliminating coexisting alternative implementations of the grammatical property.

Thus, the findings could be interpreted along the lines of the motto proposed in a paper by Chris Schaner-Wolles (1995/6) on the acquisition of negation with the title:
"From 'Anything goes' to 'Rien ne va plus.' ":
Initially, "anything goes" (crucially: within the limits of UG), until a conclusive
'rien ne va plus' guarantees the stable stage of the core grammar.

Appendix

Data Evaluation Scheme…… 138

Results of Data Evaluation (*Aleksandra*-Corpus, files 10-37):

Table 1, files *Aleks* 10-14…….. 140
Table 2, files *Aleks* 15-19…….. 142
Table 3, files *Aleks* 20-24…….. 144
Table 4, files *Aleks* 25-29…….. 146
Table 5 files *Aleks* 30-34…….. 148
Table 6, files *Aleks* 35-37…….. 152

Data Evaluation Scheme (p.1)

File: ***Aleksandra __***

Age: **Analyzable utterances:** **Utt. with verb:** **(= %)**

V.Nr.	Variable type	Variable value	- subject	+ subject
1	Finiteness: +/- finite verb vs. imperative	1: imperative 2: finite verb 3: infinitive 4: participle only (s.V.6)		
2	Modal verbs: (+/- infinitive)	1: modal verb only 2: mod./aux + infinitive 3 : aux. + participle		
3	Lexical verbs: (fin. Verbs)	1: main verb 2: copula 3: type *iść spać* (*go sleep)*		
4	Morphology finite verb present tense	1: 1st pers. sing. 2: 2nd pers. sing. 3: 3rd ´pers. sing. 4: 1st pers. plural 5: 2nd pers. plural 6 : 3rd pers. plural 7: agreement error		
5	Morphology finite verb past tense	1: 1st pers. sing. 2: 2nd pers. sing. 3: 3rd ´pers. sing. 4: 1st pers. plural 5: 2nd pers. plural 6 : 3rd pers. plural 7: agreement error 8: gender agr. error		
6	Future tense	1: aux (być) + infinitive 2: aux + participle 3: future aux. (cop.) 4 : participle only (ellipt)		

Data Evaluation Scheme (p. 2)

File: ***Aleksandra***

Age: **Analyzable utterances:** **Utt. with verb:** **(= %)**

7	**Future tense morphology auxiliary (być)**	**1: 1st pers. sing.** **2: 2nd pers. sing.** **3: 3rd ´pers. sing.** **4: 1st pers. plural** **5: 2nd pers. plural** **6 : 3rd pers. plural** **7: agreement error**		

8	**Interpretation Root Infinitive**	**1 : modal /volititional** **2: ellipt./ context-lic.** **2 : clearly non-mod./-ell.** **3 : unclear**		

9	**Lexical content (overt) subject**	**1: Noun / NP** **2: personal pronoun** **3: other pron.** (eg dem.)	
10	**Lexical content direct object (Akk.)**	**1: Noun / NP** **2: strong pronoun** **3: clitic pronoun** **4: demonstr. pronoun** **5: refl. pron (się/siebie)** **6: missing pron.** **7: wrong pronoun**	
11	**Lexical content indirect / prepositional object**	**1: Noun / NP** **2: strong pronoun** **3: clitic pronoun** **4: demonstr. pronoun** **5: reflexive pronoun (sobie, sobą)** **6: missing pron.** **7: wrong pronoun**	

V.-Nr.	Variable value # verb utterances	*Aleksandra 10* (1;09.14) # = 29			*Aleksandra 11* (1;10.0) # = 28			*Aleksandra 12* (1;10.13) # = 43			*Aleksandra 13* (1;11.09) # = 39			*Aleksandra 14* (2;0.01) # = 32		
		w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total
1	**Finiteness:**															
	1: imperative	5	-	5	2	1	3	7	-	7	8	-	8	11	-	11
	2: finite verb	14	7	**21**	16	8	**24**	21	10	**31**	24	4	**28**	15	4	**19**
	3: infinitive	2	-	2	-	-	0	4	-	4	2	-	2	1	-	1
	4: bare participle	1	-	1	1	-	1	1	-	1	1	-	1	1	-	1
		22	7	**29**	19	9	**28**	34	9	**43**	35	4	**39**	28	4	**32**
2	**Verb type: (finite verbs): modal/aux. verbs:**															
	1: modal verb only	-	-	0	-	1	1	-	-	0	1	-	1	-	-	0
	2: mod./aux.+ inf.	-	1	1	-	-	0	-	-	0	2	-	2	3	1	4
	3: aux. + partic.	-	-	0	-	-	0	-	2	2	-	-	0	-	-	0
3	**lexical verbs:**															
	1: main verb	14	7	21	15	7	22	20	5	25	19	2	21	11	2	13
	2: copula	-	-	0	-	-	0	1	3	4	2	2	4	1	1	2
	3: 2 lexical verbs	-	-	0	1	-	1	-	-	0	-	-	0	-	-	0
		14	7	**21**	16	8	**24**	21	10	**31**	24	4	**28**	15	4	**19**
4	**Verb morphology present tense:**															
	1: 1st pers. sing.	-	-	0	3	-	3	-	-	0	3	-	3	4	-	4
	2: 2nd pers. sing.	-	-	0	1	-	1	-	-	0	6	-	6	3	-	3
	3: 3rd pers. sing.	9	6	15	10	5	15	15	6	21	11	4	15	5	3	8
	4: 1st pers. plural	1	-	1	-	-	0	-	-	0	-	-	0	1	-	1
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	1	-	1	1	-	1	2	-	2	-	-	0	-	-	0
	7: agreement error	-	-	0	-	1	1	-	-	0	-	-	0	-	-	0
5	**past tense:**															
	1: 1st pers. sing.	-	-	0	-	-	0	-	-	0	1	-	1	-	-	0
	2: 2nd pers. sing.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	3: 3rd pers. sing.	2	1	3	1	2	3	3	1	4	3	-	3	2	-	2
	4: 1st pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	8: gender agr. error	-	-	0	-	1	1	-	-	0	-	-	0	-	-	0

6	**future tense:**	***Aleks 10*** **cont.**			***Aleks 11*** **cont.**			***Aleks 12*** **cont.**			***Aleks 13*** **cont.**			***Aleks 14*** **cont.**		
	1: aux. (by') + inf.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	2: aux. + participle	-	-	0	-	-	0	-	2	2	-	-	0	-	-	0
	3: future aux.(cop.)	-	-	0	-	-	0	-	2	2	-	-	0	-	-	0
	4: bare partic.(ellip)	1	-	1	1	-	1	1	-	1	1	-	1	1	-	1
7	**fut. tense morph. aux. by':**															
	1: 1st pers. sing.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	2: 2nd pers. sing.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	3: 3rd pers. sing.	-	-	0	-	-	0	-	2	2	-	-	0	-	-	0
	4: 1st pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
8	**Interpretation RI**															
	1: modal / volitional	1	-	1	-	-	0	3	-	3	2	-	2	1	-	1
	2: elliptical/cont.lic.	1	-	1	-	-	0	-	-	0	-	-	0	-	-	0
	3: **non**-mod./-ellipt.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	4: int. unclear	-	-	0	-	-	0	1	-	1	-	-	0	-	-	0
9	**Lexical content (overt) subject**							**e.g.**								
	1: Noun / NP			7			9			6			4			4
	2: pers. pronoun			0			0			0			0			0
	3: other pron (dem)			0			0	to, ten		3			0			0
10	**Lexical content direct object** (acc)	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			5			7			6			8			5
	2: strong pronoun			0			0	nas		1	ich		1			0
	3: clitic pronoun			0			0			0			0	civ		1
	4: demonstr. pron.	to		1	to		1	tego		1			0			0
	5: reflexive pr. (siv)			0			0			0	siv		1	siv		1
	6: missing pronoun	siv		1			0			0			0			0
	7: wrong pronoun			0			0			0			1			0
11	**Lex. cont. ind.obj**	**e.g.**														
	1: Noun / NP			0			1			0			4			4
	2: strong pronoun			0			0			0			0			0
	3: clitic pronoun			0			0			0	mu		1			0
	4: demonstr. pron.			0			0			0			0			0
	5: reflex. Pr. (sobie)	sobie		1			0			0			0			0

V.-Nr.	Variable value	*Aleksandra 15*			*Aleksandra 16*			*Aleksandra 17*			*Aleksandra 18*			*Aleksandra 19*		
	# verb utterances	(2;0.15) # = 89			(2;0.29) # = 143			(2;01.21) # = 52			(2;02.3) # = 153			(2;02.20) # = 125		
		w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total
1	**Finiteness:**															
	1: imperative	21	-	21	26	-	26	8	-	8	24	-	24	23	-	23
	2: finite verb	32	30	**62**	80	30	**110**	35	6	**41**	73	48	**121**	56	43	**99**
	3: infinitive	3	-	3	5	-	5	1	-	1	8	-	8	3	-	3
	4: bare participle	3	-	3	2	-	2	2	-	2	-	-	0	-	-	0
		59	30	**89**	113	30	**143**	46	6	**52**	105	48	**153**	82	43	**125**
2	**Verb type: (finite verbs): modal/aux. verbs:**															
	1: modal verb only	-	-	0	1	-	1	-	-	0	-	1	1	2	-	2
	2: mod./aux.+ inf.	1	1	2	10	1	11	5	1	6	10	3	13	6	5	11
	3: aux. + partic.	1	5	6	4	2	6	-	-	0	4	5	9	2	-	2
3	**lexical verbs:**															
	1: main verb	23	14	37	58	17	75	26	4	30	49	30	79	39	21	60
	2: copula	3	10	13	5	9	14	4	1	5	9	9	18	6	17	23
	3: 2 lexical verbs	4	-	4	2	1	3	-	-	0	1	-	1	1	-	1
		32	30	**62**	80	30	**110**	35	6	**41**	73	48	**121**	56	43	**99**
4	**Verb morphology present tense:**															
	1: 1st pers. sing.	3	-	3	16	2	18	16	-	16	15	3	18	11	11	22
	2: 2nd pers. sing.	11	3	14	10	2	12	1	2	3	11	1	12	10	2	12
	3: 3rd pers. sing.	10	12	22	31	15	47	10	4	14	19	26	45	15	18	33
	4: 1st pers. plural	2	-	2	2	-	2	1	-	1	1	-	1	-	-	0
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	-	-	0	-	-	0	3	-	3	-	2	2	5	1	6
	7: agreement error	-	-	0	1	1	2	-	-	0	-	-	0	1	-	1
5	**past tense:**															
	1: 1st pers. sing.	1	-	1	5	-	5	2	-	2	7	1	8	5	-	5
	2: 2nd pers. sing.	1	-	1	4	1	5	-	-	0	3	1	4	3	-	3
	3: 3rd pers. sing.	4	1	5	5	4	9	2	-	2	8	7	15	4	4	8
	4: 1st pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	-	-	0	1	-	1	-	-	0	-	-	0	-	-	0
	7: agreement error	-	-	0	2	-	2	-	-	0	-	-	0	-	-	0
	8: gender agr. error	-	-	0	1	-	1	-	-	0	-	-	0	-	-	0

6	**future tense:**	***Aleks 15 cont.***			***Aleks 16* cont.**			***Aleks 17* cont.**			***Aleks 18* cont.**			***Aleks 19* cont.**		
	1: aux. (by') + inf.	-	-	0	4	-	4	-	-	0	5	-	5	1	2	3
	2: aux. + participle	1	5	6	4	2	6	-	-	0	4	5	9	2	-	2
	3: future aux.(cop.)	-	5	5	-	4	4	-	-	0	-	2	2	-	5	5
	4: bare partic.(ellip)	3	-	3	2	-	2	2	-	2	-	-	0	-	-	0
7	**fut. tense morph. aux. by':**															
	1: 1st pers. sing.	-	-	0	1	1	0	-	-	0	-	-	0	-	-	0
	2: 2nd pers. sing.	-	-	0	2	2	0	-	-	0	-	-	0	-	-	0
	3: 3rd pers. sing.	1	10	11	5	2	0	-	-	0	9	7	16	3	4	7
	4: 1st pers. plural	-	-	0	-	1	0	-	-	0	-	-	0	-	-	0
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	-	1	1	-	-	0	-	-	0	-	-	0	-	3	3
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
8	**Interpretation RI**															
	1: modal / volitional	1	-	1	3	-	3	1	-	1	6	-	6	3	-	3
	2: elliptical/cont.lic.	1	-	1	2	-	2	-	-	0	2	-	2	-	-	0
	3: **non**-mod./-ellipt.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	4: int. unclear	1	-	1	-	-	0	-	-	0	-	-	0	-	-	0
9	**Lexical content (overt) subject**	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			21			23			3			34			24
	2: pers. pronoun	ty		3	ja, ty		5	ty		2	ja, ty, on		9	ja, ty, ona, one		14
	3: other pron (dem)	to		6	to		3	to		1	to		5	to, ten		5
10	**Lexical content direct object** (acc)	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			6			28			10			29			18
	2: strong pronoun			-	mnie		1			0	mnie, jego		2	mnie		1
	3: clitic pronoun	jq		2	civ		1	jq		2			0	civ, go, jq		3
	4: demonstr. pron.	to		1	to		2			0			0	to		1
	5: reflexive pr. (siv)			0			0	siv		5	siv		9	siv		13
	6: missing pronoun	siv		1	siv		1			0	siv		3	siv		3
	7: wrong pronoun			0			0			0			0			0
11	**Lex. cont. ind.obj**	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			3			3			3			18			12
	2: strong pronoun	niej,mnie	ciebie	3			0	ciebie, niego		3	mnie, ciebie,niego		8	niej, niego		3
	3: clitic pronoun	jej		1	mi, ci, mu		7			0	mi, ci		4	mi, ci, mu		6
	4: demonstr. pron.			0			0			0			0			0
	5: reflex. Pr. (sobie)			-			0			0			0			0

V.-Nr.	Variable value	*Aleksandra 20*			*Aleksandra 21*			*Aleksandra 22*			*Aleksandra 23*			*Aleksandra 24*		
	# verb utterances	(2;03.4) # = 161			(2;03.18) # = 133			(2;04.4) # = 114			(2;04.18) # = 127			(2;05.1) # = 115		
		w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total
1	**Finiteness:**															
	1: imperative	16	-	16	14	-	14	19	-	19	7	-	7	19	-	19
	2: finite verb	72	63	**135**	69	41	**110**	50	38	**88**	53	62	**115**	51	42	**93**
	3: infinitive	6	-	6	6	-	6	7	-	7	5	-	5	3	-	3
	4: bare participle	4	-	4	3	-	3	-	-	0	-	-	0	-	-	0
		98	63	**161**	92	41	**133**	76	38	**114**	65	62	**127**	73	42	**115**
2	**Verb type: (finite verbs): modal/aux. verbs:**															
	1: modal verb only	1	-	1	-	-	0	-	-	0	3	-	3	5	1	6
	2: mod./aux.+ inf.	20	3	23	6	4	10	12	5	17	6	3	9	6	4	10
	3: aux. + partic.	7	2	9	3	2	5	2	2	4	3	1	4	2	2	4
3	**lexical verbs:**															
	1: main verb	41	36	77	52	26	78	31	14	45	37	29	66	32	20	52
	2: copula	1	20	21	4	8	12	2	16	18	4	29	33	4	15	19
	3: 2 lexical verbs	2	2	4	4	1	5	3	1	4	-	-	0	2	-	2
		72	63	**135**	69	41	**110**	50	38	**88**	53	62	**115**	51	42	**93**
4	**Verb morphology present tense:**															
	1: 1st pers. sing.	17	18	35	15	6	21	10	6	16	8	5	13	15	8	23
	2: 2nd pers. sing.	7	7	14	2	1	3	5	1	6	8	4	12	4	-	4
	3: 3rd pers. sing.	17	21	38	12	14	26	9	18	27	15	43	58	10	22	32
	4: 1st pers. plural	1	2	3	3	-	3	4	-	4	1	-	1	2	-	2
	5: 2nd pers. plural	-	-	0	1	-	1	1	-	1	-	-	0	-	-	0
	6: 3rd pers. plural	2	-	2	4	2	6	4	5	9	5	3	8	-	5	5
	7: agreement error	-	1	1	1	1	2	-	-	0	-	-	0	-	1	1
5	**past tense:**															
	1: 1st pers. sing.	5	2	7	6	2	8	5	-	5	6	2	8	6	3	9
	2: 2nd pers. sing.	1	1	2	5	1	6	-	-	0	1	-	1	-	-	0
	3: 3rd pers. sing.	6	4	10	14	9	23	3	2	5	4	2	6	7	1	8
	4: 1st pers. plural	2	1	3	-	-	0	3	1	4	-	-	0	1	-	1
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	1	-	1	1	2	3	1	-	1	-	-	0	1	-	1
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	8: gender agr. error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0

6	**future tense:**	***Aleks 20 cont.***			***Aleks 21* cont.**			***Aleks 22* cont.**			***Aleks 23* cont.**			***Aleks 24* cont.**		
	1: aux. (by') + inf.	5	-	5	-	2	2	2	2	4	-	1	1	1	1	2
	2: aux. + participle	7	2	9	3	2	5	2	2	4	3	1	4	2	2	4
	3: future aux.(cop.)	1	4	5	-	-	0	1	1	2	2	-	2	2	-	2
	4: bare partic.(ellip)	4	-	4	3	-	3	-	-	0	-	-	0	-	-	0
7	**fut. tense morph. aux. by':**															
	1: 1st pers. sing.	-	-	0	1	-	1	2	-	2	3	-	3	2	1	3
	2: 2nd pers. sing.	-	-	0	1	1	2	-	1	1	-	-	0	-	-	0
	3: 3rd pers. sing.	9	6	15	1	3	4	1	4	5	2	1	3	2	2	4
	4: 1st pers. plural	2	-	2	-	-	0	-	-	0	-	-	0	1	-	1
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	2	-	2	-	-	0	2	-	2	-	1	1	-	-	0
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
8	**Interpretation RI**															
	1: modal / volitional	2	-	2	2	-	2	1	-	1	3	-	3	1	-	1
	2: elliptical/cont.lic.	3	-	3	4	-	4	6	-	6	2	-	2	2	-	2
	3: **non**-mod./-ellipt.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	4: int. unclear	1	-	1	-	-	0	-	-	0	-	-	0	-	-	0
9	**Lexical content (overt) subject**	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			24			26			19			35			8
	2: pers. pronoun	ja, ty, **my**		31	ja, ty		13	ja, ty, on		10	ja, ty, on, ona, oni		17	ja, on, ona, oni,one		24
	3: other pron (dem)	to, ten		8	to		2	to		9	to		10	to, te		10
10	**Lexical content direct object** (acc)	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			21			21			11			19			15
	2: strong pronoun	ciebie, jego		2	mnie, ciebie, nas		13	mnie		4	mnie		4	mnie, ciebie		4
	3: clitic pronoun	jq, go		3	go, jq		2	jq, je		2	civ, je		2	go, jq, ich		4
	4: demonstr. pron.	to		1	to		2	to		1	tq		1			0
	5: reflexive pr. (siv)	siv		21	siv		13	siv		17	siv		6	siv		14
	6: missing pronoun	siv		1			0			0			0			0
	7: wrong pronoun			0			0			0			0			0
11	**Lex. cont. ind.obj**	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			19			14			10			12			8
	2: strong pronoun	mnie,niej,	ciebie	7	mnie, ciebie,niego		5	mnie, niego, nas		4	niego, niej		2	jemu, niej, niego		5
	3: clitic pronoun	mi		4	mi		2	mi		4	mi, ci		4	mi, ci, mu		3
	4: demonstr. pron.			0			0			0			0			0
	5: reflex. Pr. (sobie)	siebie		1			0			0			0			0

V.-Nr.	Variable value	*Aleksandra 25*			*Aleksandra 26*			*Aleksandra 27*			*Aleksandra 28*			*Aleksandra 29*		
	# verb utterances	(2;05.15) # = 81			(2;06.0) # = 137			(2;06.16) # = 164			(2;07.0) # = 157			(2;07.14) # = 165		
		w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total
1	**Finiteness:**															
	1: imperative	5	-	5	15	-	15	24	-	24	16	-	16	16	-	16
	2: finite verb	47	25	**72**	75	42	**117**	81	52	**133**	74	63	**137**	87	59	**146**
	3: infinitive	4	-	4	4	-	4	5	-	5	4	-	4	2	-	2
	4: bare participle	-	-	0	1	-	1	2	-	2	-	-	0	1	-	1
		56	25	**81**	95	42	**137**	112	52	**164**	94	63	**157**	106	59	**165**
2	**Verb type: (finite verbs): modal/aux. verbs:**															
	1: modal verb only	3	-	3	2	-	2	2	1	3	-	2	2	-	1	1
	2: mod./aux.+ inf.	8	1	9	19	2	21	10	4	14	13	3	16	15	3	18
	3: aux. + partic.	3	1	4	-	1	1	2	1	3	3	2	5	3	3	6
3	**lexical verbs:**															
	1: main verb	31	13	44	49	27	76	62	25	87	50	40	90	66	29	95
	2: copula	2	10	12	4	11	15	5	20	25	7	16	23	2	23	25
	3: 2 lexical verbs	-	-	0	1	1	2	-	1	1	1	-	1	1	-	1
		47	25	**72**	75	42	**117**	81	52	**133**	74	63	**137**	87	59	**146**
4	**Verb morphology present tense:**															
	1: 1st pers. sing.	9	2	11	27	5	32	29	7	36	16	17	33	25	6	31
	2: 2nd pers. sing.	5	-	5	4	-	4	12	3	15	8	6	14	-	2	2
	3: 3rd pers. sing.	4	14	18	14	16	30	15	26	41	7	14	21	11	27	38
	4: 1st pers. plural	12	-	12	3	-	3	2	-	2	5	-	5	11	1	12
	5: 2nd pers. plural	1	-	1	-	-	0	-	-	0	1	-	1	-	-	0
	6: 3rd pers. plural	1	1	2	3	3	6	3	4	7	6	4	10	-	2	2
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
5	**past tense:**															
	1: 1st pers. sing.	1	1	2	10	5	15	7	-	7	11	7	18	6	2	8
	2: 2nd pers. sing.	2	-	2	5	2	7	2	1	3	2	3	5	1	-	1
	3: 3rd pers. sing.	5	3	8	6	8	14	5	7	12	10	9	19	23	10	33
	4: 1st pers. plural	-	-	0	-	-	0	-	-	0	1	-	1	-	-	0
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	1	-	1	-	-	0
	6: 3rd pers. plural	-	-	0	-	1	1	1	1	2	-	-	0	1	1	2
	7: agreement error	-	-	0	-	-	0	1	-	1	-	-	0	-	-	0
	8: gender agr. error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0

6	**future tense:**	***Aleks 25 cont.***			***Aleks 26* cont.**			***Aleks 27* cont.**			***Aleks 28* cont.**			***Aleks 29* cont.**		
	1: aux. (by') + inf.	2	-	2	4	1	5	1	-	1	1	1	2	5	1	6
	2: aux. + participle	3	1	4	-	1	1	2	1	3	3	2	5	2	3	5
	3: future aux.(cop.)	1	3	4	1	-	1	-	1	1	2	-	2	1	4	5
	4: bare partic.(ellip)	-	-	0	1	-	1	2	-	2	-	-	0	1	-	1
7	**fut. tense morph. aux. by':**															
	1: 1st pers. sing.	3	-	3	1	-	1	2	-	2	3	1	4	4	4	8
	2: 2nd pers. sing.	1	-	1	1	-	1	-	-	0	-	-	0	-	-	0
	3: 3rd pers. sing.	-	3	3	1	2	3	1	2	3	1	2	3	1	4	5
	4: 1st pers. plural	-	-	0	1	-	1	-	-	0	-	-	0	3	-	3
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	2	1	3	1	-	1	-	-	0	2	-	2	-	-	0
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
8	**Interpretation RI**															
	1: modal / volitional	1	-	1	3	-	3	3	-	3	2	-	2	2	-	2
	2: elliptical/cont.lic.	3	-	3	1	-	1	2	-	2	2	-	2	-	-	0
	3: **non**-mod./-ellipt.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	4: int. unclear	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
9	**Lexical content (overt) subject**	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			13			24			26			25			30
	2: pers. pronoun	ja, on, ona		6	ja, ty, ona		14	ja, ty, on, ona		18	ja, ty, on, ona,one		32	ja, ty, on, ona, my		21
	3: other pron (dem)	to		6	to		4	to		8	to, ten		6	to		8
10	**Lexical content direct object** (acc)	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			12			24			11			31			32
	2: strong pronoun	mnie		1	mnie, ciebie, jego		7	mnie, ciebie		9	mnie, ciebie, nas		14	ciebie		1
	3: clitic pronoun	jq, go		2	jq		2	jq, civ, go		4	civ, go		3	jq		1
	4: demonstr. pron.			0			0			0	to, te		3	tq		1
	5: reflexive pr. (siv)	siv		6	siv		14	siv		5	siv		9	siv		19
	6: missing pronoun	siv		1			0	siv		1	siv		1			0
	7: wrong pronoun			0			0			0			0			0
11	**Lex. cont. ind.obj**	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			9			14			14			11			13
	2: strong pronoun	jemu, nich, nimi		3	ciebie, niego		3	niego, niej, nich		9	jej, niego, niej		15	niej, niego, nam		3
	3: clitic pronoun	mi		1	mi		3	mi, ci, mu		8	mi		5	mi, ci		5
	4: demonstr. pron.			0			0			0			0			0
	5: reflex. Pr. (sobie)			0			0			0	sobie		4	sobie		7

V.-Nr.	Variable value	*Aleksandra 30*			*Aleksandra 31*			*Aleksandra 32*			*Aleksandra 33*			*Aleksandra 34*		
	# verb utterances	(2;08.0) # = 135			(2;08.15) # = 126			(2;09.7) # = 149			(2;10.1) # = 132			(2;11.7) # = 123		
		w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total
1	**Finiteness:**															
	1: imperative	9	-	9	17	-	17	21	-	21	21	-	21	12	-	12
	2: finite verb	68	55	**123**	60	46	**106**	64	59	**123**	65	44	**109**	63	47	**110**
	3: infinitive	3	-	3	3	-	3	3	-	3	2	-	2	1	-	1
	4: bare participle	-	-	0	-	-	0	2	-	2	-	-	0	-	-	0
		80	55	**135**	80	46	**126**	90	59	**149**	88	44	**132**	76	47	**123**
2	**Verb type: (finite verbs): modal/aux. verbs:**															
	1: modal verb only	1	1	2	-	-	0	1	2	3	2	-	2	1	-	1
	2: mod./aux.+ inf.	12	2	14	9	6	15	17	6	23	10	4	14	10	3	13
	3: aux. + partic.	2	4	6	3	-	3	-	2	2	-	2	2	4	-	4
3	**lexical verbs:**															
	1: main verb	48	27	75	42	18	60	41	16	57	49	18	67	38	23	61
	2: copula	5	20	25	6	22	28	5	33	38	2	20	22	10	21	31
	3: 2 lexical verbs	-	1	1	-	-	0	-	-	0	2	-	2	-	-	0
		68	55	**123**	60	46	**106**	64	59	**123**	65	44	**109**	63	47	**110**
4	**Verb morphology present tense:**															
	1: 1st pers. sing.	27	13	40	19	6	25	27	8	35	17	9	26	14	9	23
	2: 2nd pers. sing.	3	4	7	6	1	7	5	1	6	19	1	20	11	2	13
	3: 3rd pers. sing.	10	10	20	16	33	49	9	18	27	11	22	33	11	13	24
	4: 1st pers. plural	2	-	2	1	-	1	3	-	3	-	-	0	1	-	1
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	1	-	1
	6: 3rd pers. plural	6	5	11	3	-	3	1	3	4	2	2	4	2	4	6
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
5	**past tense:**															
	1: 1st pers. sing.	9	3	12	3	1	4	7	-	7	4	1	5	4	1	5
	2: 2nd pers. sing.	3	1	4	4	-	4	-	-	0	3	-	3	1	1	2
	3: 3rd pers. sing.	5	9	14	1	4	5	7	5	12	5	6	11	10	11	21
	4: 1st pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	-	1	1	-	1	1	-	2	2	-	-	0	1	3	4
	7: agreement error	-	-	0	-	-	0	1	-	1	-	-	0	-	-	0
	8: gender agr. error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0

6	future tense:	*Aleks 30 cont.*			*Aleks 31* cont.			*Aleks 32* cont.			*Aleks 33* cont.			*Aleks 34* cont.		
	1: aux. (by') + inf.	1	1	2	1	-	1	3	4	7	2	-	2	-	-	0
	2: aux. + participle	2	4	6	3	-	3	-	2	2	-	2	2	4	-	4
	3: future aux.(cop.)	0	3	3	1	-	1	2	16	18	-	1	1	2	3	5
	4: bare partic.(ellip)	-	-	0	-	-	0	1	-	1	-	-	0	-	-	0
7	**fut. tense morph. aux. by':**															
	1: 1st pers. sing.	1	2	3	2	-	2	2	3	5	1	-	1	3	-	3
	2: 2nd pers. sing.	-	1	1	1	-	1	-	2	2	-	1	1	1	-	1
	3: 3rd pers. sing.	2	5	7	2	-	2	1	13	14	-	2	2	2	3	5
	4: 1st pers. plural	-	-	0	-	-	0	1	-	0	1	-	1	-	-	0
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	-	-	0	-	-	0	1	4	5	-	-	0	-	-	0
	7: agreement error	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
8	**Interpretation RI**															
	1: modal / volitional	1	-	1	2	-	2	-	-	0	1	-	1	-	-	0
	2: elliptical/cont.lic.	2	-	2	1	-	1	3	-	3	1	-	1	1	-	1
	3: **non**-mod./-ellipt.	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
	4: int. unclear	-	-	0	-	-	0	-	-	0	-	-	0	-	-	0
9	**Lexical content (overt) subject**	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			28			16			28			12			16
	2: pers. pronoun	ja, ty, on, ona, oni		24	ja, ty, on, ona		16	ja, ty, on,ona, one		22	ja, ty, on, ona		20	ja, ty, on, ona, oni		23
	3: other pron (dem)	to, te		3	to, ten		14	to, ten, ta		9	to		11	to, ten		8
10	**Lexical content direct object** (acc)	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			19			25			15			19			17
	2: strong pronoun	mnie, ciebie		2			0	mnie		1	mnie, ciebie		3	mnie		3
	3: clitic pronoun	je, civ		2	civ, je		2	jq, go		4	civ, go, jq		7	jq, je		4
	4: demonstr. pron.	to, tego		2	to		1			0	to, tego		5	tq		1
	5: reflexive pr. (siv)	siv		13	siv		7	siv		12	siv		12	siv		7
	6: missing pronoun	siv		1			0			0			0	siv		3
	7: wrong pronoun			0			0			0			0			0
11	**Lex. cont. ind.obj**	**e.g.**			**e.g.**			**e.g.**			**e.g.**			**e.g.**		
	1: Noun / NP			12			7			8			6			7
	2: strong pronoun	jemu, nich, nami		10	mnie, ciebie,niego		5	mnie, ciebie, nich		9	mnie, ciebie, jemu		3	mnie ,niej, nich		4
	3: clitic pronoun	mi, ci		7	mi		3	mi, mu		2	mi, ci, mu		9	mi, ci		6
	4: demonstr. pron.			0			0			0			0			0
	5: reflex. Pr. (sobie)	sobie		3	sobie		2	sobie		2	sobie		4	sobie		2

V.-Nr.	Variable value	*Aleksandra 35*			*Aleksandra 36*			*Aleksandra 37*		
	# verb utterances	(3;0.19) # = 114			(3;01.24) # = 179			(3.03.2) # = 196		
		w/o subj	w subj.	total	w/o subj	w subj.	total	w/o subj	w subj.	total
1	**Finiteness:**									
	1: imperative	7	-	7	20	-	20	17	-	17
	2: finite verb	72	31	**103**	94	60	**154**	112	65	**177**
	3: infinitive	4	-	4	4	-	4	2	-	2
	4: bare participle	-	-	0	1	-	1	-	-	0
		83	31	**114**	119	60	**179**	131	65	**196**
2	**Verb type: (finite verbs): modal/aux. verbs:**									
	1: modal verb only	-	-	0	2	-	2	1	-	1
	2: mod./aux.+ inf.	4	-	4	6	7	13	8	10	18
	3: aux. + partic.	1	-	1	4	4	8	-	-	0
3	**lexical verbs:**									
	1: main verb	58	26	84	68	37	105	78	32	110
	2: copula	9	5	14	12	12	24	23	23	46
	3: 2 lexical verbs	-	-	0	2	-	2	2	-	2
		72	31	**103**	94	60	**154**	112	65	**177**
4	**Verb morphology present tense:**									
	1: 1st pers. sing.	18	6	24	25	17	42	40	14	54
	2: 2nd pers. sing.	5	3	8	9	2	11	10	5	15
	3: 3rd pers. sing.	13	13	26	10	12	22	15	19	34
	4: 1st pers. plural	2	1	3	11	2	13	6	2	8
	5: 2nd pers. plural	1	-	1	1	-	1	1	-	1
	6: 3rd pers. plural	6	3	9	5	5	10	5	8	13
	7: agreement error	-	-	0	-	-	0	1	-	1
5	**past tense:**									
	1: 1st pers. sing.	3	1	4	5	3	8	10	4	14
	2: 2nd pers. sing.	1	-	1	-	-	0	3	-	3
	3: 3rd pers. sing.	18	4	22	18	9	27	14	4	18
	4: 1st pers. plural	-	-	0	-	-	0	1	-	1
	5: 2nd pers. plural	-	-	0	-	-	0	-	-	0
	6: 3rd pers. plural	3	1	4	1	3	4	3	-	3
	7: agreement error	-	-	0	-	-	0	-	-	0
	8: gender agr. error	-	-	0	-	-	0	-	-	0

6	**future tense:**	***Aleks 35 cont.***			***Aleks 36* cont.**			***Aleks 37* cont.**			
	1: aux. (by') + inf.	-	-	0	2	2	4	1	4	5	
	2: aux. + participle	1	-	1	4	4	8	-	-	0	
	3: future aux.(cop.)	-	-	0	2	1	3	3	4	7	
	4: bare partic.(ellip)	-	-	0	1	-	1	-	-	0	
7	**fut. tense morph. aux. by':**										
	1: 1st pers. sing.	-	-	0	1	3	4	1	1	2	
	2: 2nd pers. sing.	-	-	0	-	-	0	1	-	1	
	3: 3rd pers. sing.	1	-	1	5	3	8	2	5	7	
	4: 1st pers. plural	-	-	0	-	-	0	-	-	0	
	5: 2nd pers. plural	-	-	0	-	1	1	-	-	0	
	6: 3rd pers. plural	-	-	0	2	-	2	-	2	2	
	7: agreement error	-	-	0	-	-	0	-	-	0	
8	**Interpretation RI**										
	1: modal / volitional	1	-	1	-	-	0	-	-	0	
	2: elliptical/cont.lic.	3	-	3	4	-	4	2	-	2	
	3: **non**-mod./-ellipt.	-	-	0	-	-	0	-	-	0	
	4: int. unclear	-	-	0	-	-	0	-	-	0	
9	**Lexical content (overt) subject**	**e.g.**			**e.g.**			**e.g.**			
	1: Noun / NP			14			25			27	
	2: pers. pronoun	ja, ty, on,ona, my		17	ja, ty, on, my, wy		34	ja, ty, on, my, oni		32	
	3: other pron (dem)			0	to		1	to, ten		6	
10	**Lexical content direct object** (acc)	**e.g.**			**e.g.**			**e.g.**			
	1: Noun / NP			21			20			26	
	2: strong pronoun	mnie, ich		5	mnie, ciebie, nas		8	mnie, ich		5	
	3: clitic pronoun	go		1	civ, go, je		7	jq, go		6	
	4: demonstr. pron.			0	tego		1			0	
	5: reflexive pr. (siv)	siv		10	siv		21	siv		11	
	6: missing pronoun	siv		1			0			0	
	7: wrong pronoun			0			0			0	
11	**Lex. cont. ind.obj**	**e.g.**			**e.g.**			**e.g.**			
	1: Noun / NP			9			13			9	
	2: strong pronoun	mnie, nas		2	mnie, niego, nich		3	mnie, ciebie, jej		6	
	3: clitic pronoun	mi, mu		3	mi, ci, mu, im		11	ci		4	
	4: demonstr. pron.			0			0			0	
	5: reflex. Pr. (sobie)	sobie		2	sobie		5	sobie		4	

References

Allen, S. E. M. (1996). *Aspects of Argument Structure Acquisition in Inuktitut.* Amsterdam: John Benjamins.

Allen, S. E. M. (1995). "Learning Noun Incorporation in Early Inuktitut." Talk given at the University of Stuttgart, November 23, 1995.

Atkinson, M. (1992). *Children's Syntax. An Introduction to Principles and Parameters Theory.* Oxford: Blackwell.

Babyonyshev, M. (1994). "The Acquisition of Russian Case." *MIT Working Papers in Linguistics* 19, p. 1-45.

Bar-Shalom, E., W. Snyder & J. Boro (1996). "Evidence for the Optional Infinitive Stage in Russian." In: A. Halbert & K. Matsuoka (eds.). *Papers in Honor of Steqhen Crain: Language Acquisition and Processing.* Storrs, CT: University of Connecticut Working Papers in Linguistics.

Bar-Shalom, E. & W. Snyder (1998). "Root Infinitives in Child Russian: A comparison with Italian and Polish." In: R. Shillcock, A. Sorace & C. Heycock (eds.). *Proceedings of the GALA 1997 Conference on Language Acquisition: Knowledge Representation and Processing.*

Bar-Shalom, E. & W. Snyder (2001). "Descriptive Imperatives in Child Russian and Early Correct Use of Verbal Morphology." In: *Proceedings of the 25th Boston University Conference on Language Development (BUCLD 25),* Somerville, MA: Cascadilla Press, p. 94-101.

Becker, M. (2000a). The Development of the Copula in Child English: The Lightness of Be. PhD Dissertation, UCLA.

Becker, M. (2000b). "The Acquisition of the Copula in Child English." In: R. Billerey & D. Lillehaugen (eds.) *WCCFL 19 Proceedings.* Somerville, MA: Cascadilla Press, p. 57-70.

Behrens, H. (1993). *Temporal Reference in German Child Language. Form and Function of Early Verb Use.* Doctoral dissertation, Universiteit van Amsterdam.

Berwick, R. (1985). *The Acquisition of Syntactic Knowledge.* Cambridge, Mass.: MIT Press.

Bielec, D. (1998). *Polish: An Essential Grammar.* London: Routledge.

Blom, E. & F. Wijnen (1999). "How Dutch Children's Root Infinitives Become Modal." Talk given at the 24th Boston University Conference on Language Development.

Borer, H. & K. Wexler (1987). "The Maturation of Syntax". In: T. Roeper & E. Williams (eds.). *Parameter Setting*. Dordrecht: Kluwer, p. 123-172..

Borer, H. & K. Wexler (1992). "Biunique Relations and the Maturation of Grammatical Principles." *Natural Language and Linguistic Theory 10,* p. 1-43.

Borsley, R. & M. Rivero (1994). "Clitic Auxiliaries and Incorporation in Polish." *Natural Language and Linguistic Theory.*

Boser, K., Lust, B., Santelmann, L. & J. Whitman (1992). "The Syntax of CP and V2 in early child German (ECG): The Strong Continuity Hypothesis." *Proceedings of NELS*, 22.

Bromberg, H.S. & K. Wexler (1995). "Null Subjects in Wh-Questions." In: *MIT Working Papers in Linguistics* 26, p. 221-248.

Brown, R. (1973). *A First Language.* Cambridge, Mass.: Harvard University Press.

Brun, D., S. Avrutin & M. Babyonyshev (1999). "Aspect and Its Temporal Interpretation during the Optional Infinitive Stage in Russian." *Proceedings of the 23rd Boston University Conference on Language Development (BUCLD 23).* Somerville, MA: Cascadilla Press, p. 120-131.

Cardinaletti, A. & M. T. Guasti (1995). "Small Clauses: Some Controversies and Issues of Acquisition." In: A. Cardinaletti & M. T. Guasti (eds.). *Small Clauses.* Syntax & Semantics 28. San Diego, Calif.: Academic Press.

Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of Language.* New York: Praeger.

Chomsky, N. (1991). "Some Notes on Economy of Derivation and Representation". In: R. Freidin (ed.). *Principles and Parameters in Comparative Grammar.* Cambridge, MA: MIT Press.

Chomsky, N. (1992). "A Minimalist Program for Linguistic Theory". *MIT Occasional Papers in Linguistics 1*.

Clahsen, H. (1986). *Die Profilanalyse.* berlin: Marhold.

Clahsen, H. (1992). "Learnability Theory and the Problem of Development in Language Acquisition." In: J. Weissenborn, H. Goodluck & T. Roeper (eds.), p. 53-76.

Clahsen, H. & M. Penke (1992). "The Acquisition of Agreement Morphology and its Syntactic Consequences." In: J. Meisel (ed.). *The Acquisition of Verb Placement.* Dordrecht: Kluwer, p. 181-224.

Clahsen, H., Eisenbeiß, S. & M. Penke (1994). "Underspecification and lexical learning in Early Child Grammars." *Essex Research Reports in Linguistics* 4:1994.

Clark, R. & I. G. Roberts (1993). "A Computational Model of Language Learnability and Language Change." *Linguistic Inquiry* 24, p. 299-345.

Crago, M. & S. Allen (1994).”Building the Case for Impairment in Linguistic Representation.” In: M. Rice (ed.). *Towards a Genetics of Language.* Hillsdale, N.J.: Erlbaum.

Crisma, P. (1992). "On the Acquisition of Wh-Questions in French." *Geneva Generative Papers* , p. 115-122.

Czinglar, C., A. Katičić, K. Köhler & C. Schaner-Wolles (2003). “Strategies in the L1-Acquisition of Predication: The Copula Construction in German and Croatian.” Ms., Universität Wien.

Déprez, V. & A. Pierce (1993). "Negation and Functional Projections in Early Grammar." *Linguistic Inquiry* 24, p. 25-68.

Dressler, W. U. & Makovec-Cerne (1995). *Frühstadien der Verbalflexion bei einem slowenischen Kleinkind.* Sonderdruck aus dem Anzeiger der phil-hist. Klasse der Österreichischen Akademie der Wissenschaften. Wien 1995, 131: p. 59-66.

Eisenbeiß, S. (1994a). "Elizitation von Nominalphrasen und Kasusmarkierungen". *Theorie des Lexikons.* Arbeiten des Sonderforschungsbereichs 282 Nr. 57.

Eisenbeiß, S. (1994b). "Kasus und Wortstellungsvariation im deutschen Mittelfeld. Theoretische Überlegungen und Untersuchungen zum Erstspracherwerb." In: B. Haftka (ed.). *Was determiniert Wortstellungsvariation?* Opladen: Westdeutscher Verlag., p. 277-298.

Eisenbeiß, S. (1995). "Geben und Nehmen. Zum Verhältnis von Grammatiktheorie und Spracherwerbsforschung." Vortrag anläßlich des 13. Hamburger Kognitions-kolloquiums. Universität Hamburg, 4.2.1995.

Feldstein, R. & S. Franks (2002). *Polish*. München: Lincom Europa.

Fisiak, J. , Lipinska-Grzegorek, M. & t. Zabrocki (1978). *An Introductory Polish-English Contrastive Grammar.* Warszawa: Panstwowe Wydawnictwo Naukowe.

Fletcher, P. & B. MacWhinney (eds.) (1995). *The Handbook of Child Language.* Oxford: Basil Blackwell.

Franks, S. (1995). *Parameters of Slavic Morphosyntax.* Oxford: Oxford University Press.

Friedemann, M.-A. (1992). "The Underlying Position of External Arguments in French: A Study in Adult and Child Grammar." *Geneva Generative Papers* 1-2, p. 123- 144.

Goodluck, H. (1991). *Language Acquisition: A Linguistic Introduction.* Oxford: Blackwell.

Guasti, M. T. (1993/4). "Verb Syntax in Italian Child Grammar: Finite and Nonfinite Verbs." *Language Acquisition* 3 (1), p. 1-40.

Guasti, M. T. (2002). *Language Acquisition: The Growth of Grammar.* Cambridge, Mass.: The MIT Press.

Haag-Merz, C. (1995). Pronomen im Schwäbischen – Syntax und Erwerb. Doktorarbeit Universität Stuttgart.

Haegeman, L. (1994). "Root Infinitives, Tense and Truncated Structure." *Geneva Generative Papers* Vol. 2, No. 1.

Haider, H. (1993). *Deutsche Syntax - Generativ.* Tübingen: Narr.

Hamann, C. (1994). "Negation, Infinitives and Heads". Ms. Universität Genf.

Harris, A. & K. Wexler (1996). "The Optional Infinitive Stage in Child English: Evidence from Negation." In: H. Clahsen (ed.). *Studies in the Acquisition of Inflection.*

Hoekstra, T. & B. Schwartz (eds.) (1994). *Language Acquisition Studies in Generative Grammar.* Amsterdam: John Benjamins.

Hoekstra, T. & N. Hyams (1998). "Aspects of Root Infinitives". *Lingua* 106, p. 81-112.

Hoekstra, T. & N. Hyams (1999). "The Eventivity Constraint and Modal Reference Effects in Root Infinitives." In: *Proceedings of the 22nd Boston University Conference on Language Development (BUCLD 22).* Somerville, MA: Cascadilla Press, p. 240-252.

Hornstein, N. & D. Lightfoot (eds.) (1994). *Verb Movement.* Cambridge: Cambridge University Press.

Hyams, N. (1986). *Language Acquisition and the Theory of Parameters.* Dordrecht: Reidel.

Hyams, N. (1992). "A Reanalysis of Null Subjects in Child Language." In: J. Weissenborn, H. Goodluck, and T. Roeper (eds.). *Theoretical Issues in Language Acquisition.* Hillsdale: N.J.: Lawrence Erlbaum.

Hyams, N. (1996). "The Underspecification of Functional Categories in Early Grammar." In: H. Clahsen (ed). *Generative Perspectives on Language Acquisition.* Amsterdam : John Benjamins.

Hyams, N. & K. Wexler (1993). "On the Grammatical Basis of Null Subjects in Child Language." *Linguistic Inquiry* 24, p. 421-459.

Hyams, N. (2001). "Now You Hear It, Now You Don't: The Nature of Optionality in Child Grammars." In: *Proceedings of the 25th Boston University Conference on Language Development (BUCLD 25),* Somerville, MA: Cascadilla Press, p. 34-58.

Ingram, D. & W. Thompson (1996). "Early Syntactic Acquisition in German: vidence for the Modal Hypothesis." *Language* 72, p. 97-120.

Jonas, D. (1995). "On the Acquisition of Verb Syntax in Child Faroese." *MIT Working Papers in Linguistics* 26, p.325-382.

Jordens, P. (1990). "The Acquisition of Verb Placement in Dutch and German." *Linguistics* 28, p. 1407-1448.

Kampen, J. van & A. Evers (1995). "LF Relevance in Language Acquisition." *Amsterdam Series In Child Language Development,* 4, p. 1-31.

Kampen, J. van (1997). *First Steps in WH-Movement.* Delft: Eburon.

Katičić, A. (1997). Zum Erstspracherwerb des Kroatischen: Morphologische und syntaktische Aspekte beim Erwerb des Verbalsystems. Diplomarbeit Universität Wien.

Katičić, A. & C. Schaner-Wolles (2001). „Subjekt i predikat u usvajanju hrvatskoga jezika." Paper presented at *Psycholinguistics and Cognitive Science in Croatian Applied Linguistics. 15. Jahrestagung der Croatian Applied Society,* Opatija, May 2001.

Klepper-Schudlo, A. (1996). "Early Verbal Inflection in Polish and the Optional Infinitive Phenomenon." Paper presented at the *Acquisition of Morphology*-Workshop, *Seventh International Morphology Meeting,* Vienna.

Klepper-Schudlo, A. (1999). "Zum L1-Erwerb von Verbalflexion und Pronomina im Polnischen." Paper presented at the 27th *Österreichische Linguistiktagung*, Wien: November 1999.

Köhler, K. & S. Bruyère (1996). Finiteness and Verb Placement in the L1 acquisition of German. *Wiener Linguistische Gazette* 53-54 / 1995-96, Institut für Sprachwissenschaft der Universität Wien, p. 63-86.

Köhler, K. (1998). Finitheit und Verbposition: Eine Fallstudie zum Erstspracherwerb des Deutschen. Diplomarbeit Universität Wien.

Krämer, I. (1993). "The Licensing of Subjects in Early Child Language." *MIT Working Papers in Linguistics* 19, p. 197-212. (Papers on Case and Agreement).

Lasser, I. (1997). "The Interpretation of Root Infinitive Constructions in Adult and Child German." *MIT Occasional Papers in Linguistics* 12, p. 26-64.

Lebeaux, D. (1988). *Language Acquisition and the Form of Grammar.* PhD. Dissertation, University of Massachusetts.

LeBlanc, D. (1995). "An Activation Model of Parameter Setting." *Amsterdam Series In Child Language Development,* 4, p. 81-106.

Legendre, G., P. Hagstrom, A. Vainikka & M. Todorova (2000). "Evidence for Syntactic Competition During Acquisition of Tense and Agreement in Child French." In: *Proceedings of the Chicago Linguistics Society* 36.

MacWhinney, B. & C. Snow (1990). "The Child Language Data Exchange System: an Update." In: *Journal of Child Language* 17, p. 457-472.

MacWhinney, B. (1995). "Computational Analysis of Interactions." In: *The Handbook of Child Language*, ed. by P. Fletcher & B. MacWhinney.

Meisel, J. (ed.) (1992). *The Acquisition of Verb Placement.* Dordrecht: Kluwer.

Meisel, J. (1995). "Parameters in Acquisition". In: *The Handbook of Child Language*, p. 10-35.

Meisel, J. & N. Müller (1992). "Finiteness and Verb Placement in Early Child Grammar: Evidence from Simultaneous Acquisition of French and Germans in Bilinguals." In: J. Meisel (ed.).

Miodunka, W. & J Wróbel (1986). *Polska po polsku. Tom II: Podstawowe wiadomości gramatyczne i ćwiczenia strukturalne*. [Uniwersytet Jagielloński w Krakowie.] Warszawa: Wydawnictwo Interpress.

Moro, A. (1995). "Small Clauses with Predicative Nominals." In: A. Cardinaletti & M. T. Guasti (eds.). *Small Clauses.* Syntax and Semantics 28. San Diego, CA: Academic Press.

Moro, A.(1997). *The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structures.* Cambridge: Cambridge University Press.

Phillips, C. (1995). Syntax at Age Two: Cross-Linguistic Differences. In: *MIT Working Papers in Linguistics* 26, p. 335-382.

Phillips, C. (1996). "Root Infinitives are Finite." *Proceedings of the 22th Boston University Conference on Language Development (BUCLD 20).* Somerville, MA: Cascadilla Press.

Pierce, A. (1992). *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars.* Dordrecht: Kluwer.

Platzack, C. (1990). "A Grammar without Functional Categories. A Syntactic Study of Early Swedish Child Language." *Nordic Journal of Linguistics* 13 (2), p. 107-126.

Platzack, C. & A. Holmberg (1989)."The Role of AGR and Finiteness." *Working Papers in Scandinavian Syntax 43,* p. 51-76.

Poeppel, D. & K. Wexler (1993). "The Full Competence Hypothesis of Clause Structure in Early German." In: *Language 69,* p. 1-33.

Pollock, J.-Y. (1989). "Verb Movement, Universal Grammar and the Structure of IP." *Linguistic Inquiry* 20, p. 365-424.

Radford, A. (1990). *Syntactic Theory and the Acquisition of English Syntax.* Oxford: Basil Blackwell.

Radford, A. (1994). "The Syntax of Questions in Child English." *Journal of Child Language* 21, p. 211-236.

Radford, A. (1996). "Toward a structure-Building Model of Language Acquisition." In: H. Clahsen (ed). *Generative Perspectives on Language Acquisition.* Amsterdam: John Benjamins.

Radford, A. (1997). *Syntactic Theory and the Structure of English.* Cambridge: Cambridge University Press.

Rhee, J. & K. Wexler (1995). "Optional Infinitives in Hebrew." In: *MIT Working Papers in Linguistics* 26, p. 383-402.

Rizzi, L. (1990). *Relativized Minimality*. MIT Press.

Rizzi, L. (1992). "Early Null Subjects and Root Null Subjects." *Geneva Generative Papers*, p. 102-114.

Rizzi, L. (1994). "Some Notes on Linguistic Theory and Language Development. The Case of Root Infinitives." *Geneva Generative Papers*, p. 16-25.

Rizzi, L. (1998). "Remarks on Early Null Subjects." Ms.

Roeper, T. (1992). "From the Initial State to V2: Acquisition Principles in Action." In: J. Meisel (ed.).

Roberts, I. G. (1994). "Universal Grammar and L1 Acquisition." Ms., University of Wales Bangor.

Roberts, I. G. (1997). *Comparative Syntax.* London: Arnold.

Roberts, I. G. (1999). "Verb Movement and Markedness." In: M. DeGraff (ed.) (1999). *Language Creation and Language Change: Creolization, Diachrony, and Development.* Cambridge, Mass.: The MIT Press.

Rothweiler, M. (ed.) (1990). *Spracherwerb und Grammatik:Linguistische Überlegungen zum Erwerb von Syntax und Morphologie.* Opladen: Westdeutscher Verlag. [= Linguistische Berichte, Sonderheft 3]

Saloni, Z. & M. Świdzinski (1985). *Składnia współczesnego języka polskiego.* Warszawa: Państwowe Wydawnictwo Naukowe.

Schönenberger, M., A. Pierce, K. Wexler & F. Wijnen (1995). Accounts of Root Infinitives and the Interpretation of Root Infinitives. *Geneva Generative Papers*, p. 47-71.

Schaner-Wolles, C. (1996). The Acquisition of Negation in a Verb-Second Language. In: *Wiener Linguistische Gazette* 53-54 / 1995-96, p. 87-119.

Schaner-Wolles, C. (2000). "Within-Language Dissociations in Mental Retardation: Williams-Beuren and Down Syndrome." *Proceedings of the 24th Boston University Conference on Language Development (BUCLD 24).* Somerville, MA: Cascadilla Press.

Schaner-Wolles, C. (2000). „Der Erwerb von Finitheit." Paper presented at *Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft.* Universität Marburg, März 2000.

Slobin, D. I. (ed.) (1985). *The Crosslinguistic Study of Language Acquisition.* Vol.2: *Theoretical Issues.* Hillsdale, NJ: Erlbaum.

Smoczynska, M. (1985). "The Acquisition of Polish." In: D. I. Slobin (ed.). *The Crosslinguistic Study of Language Acquisition.* Vol.2: *Theoretical Issues.* Hillsdale, NJ: Erlbaum. p. 595-675.

Snyder, W. & E. Bar-Shalom (1998). "Word Order, Finiteness, and Negation in Early Child Russian." *Proceedings of the 22nd Boston University Conference on Language Development (BUCLD 22).* Somerville, MA: Cascadilla Press.

Stowell, T. (1983). “Subjects across categories.” *The Linguistic Review* 2, p. 285-312.

Szczegielniak, A. (1993). "Cliticization in Polish." Master Thesis, University of Cambridge, Clare College, Great Britain.

Tajsner, P. (1990). *Case Marking in English and Polish: A Government and Binding Study.* Doctoral dissertation, Adam Mickiewicz University, Poznan.

Tajsner, P. (1993). "Some Implications of the Minimalist Program for the Analysis of Polish Word Order." Talk at the "XXVIII International Conference on Cross-Language Studies and Contrastive Linguistics", Rydzyna (Poland), 9.-11.12.1993.

Tajsner, P. (1995). "Licensing of bare NP adjuncts." Unpublished Ms., Adam Mickiewicz University, Poznań..

Torrens, V. (1995). “The Acquisition of Inflection in Spanish and Catalan.” In: C. T. Schütze, J. B. Ganger & K. Broihier (eds.). *Papers on Language Processing and Acquisition.* MIT Working Papers in Linguistics 26. Cambridge, Mass.: MIT, MITWPL.

Tracy, R. (1990). "Spracherwerb trotz Input." In: M. Rothweiler (ed.), p. 22-49.

Tracy, R. & E. Lattey (eds.) (1994). *How Tolerant is Universal Grammar? Essays on Language Learnability and Language Variation.* Tübingen: Niemeyer.

Valian, V. (1991). “Syntactic Subjects in the Early Speech of American and Italian Children.” *Cognition* 40, p. 21-81.

Valian, V. & Z. Eisenberg (1996). “The Development of Syntactic Subjects in Portuguese-Speaking Children. *Journal of Child Language* 23, p. 103-128.

Varlokosta, S., A. Vainikka & B. Rohrbacher (1996). “Root Infinitive without Infinitives.” In: *Proceedings of the 22th Boston University Conference on Language Development (BUCLD 20).* Somerville, MA: Cascadilla Press, p. 816-827.

Verrips, M. & J. Weissenborn (1992). "Routes to Verb Placement in Early German and French: the Independence of Finiteness and Agreement." In: J. Meisel (ed.). Weissenborn, J., H. Goodluck & T. Roeper (eds.) (1992). *Theoretical Issues in Language Acquisition: Continuity and Change in Development.* Hillsdale, NJ: Erlbaum.

Verrips, M. (1994). “Learnability Meets Development: The Case of Pro-Drop.In. R. Tracy & E. Lattey (eds.). *How tolerant is Universal Grammar? Essays on Language Learnability and Language Variation. Tübigen: Niemeyer.*

Vikner, S. (1995). *Verb Movement and Expletive Subjects in the Germanic Languages.* Oxford: Oxford University Press.

Weissenborn, J. (1990). "Functional Categories and Verb Movement: The Acquisition of German Syntax Reconsidered." In: M. Rothweiler (ed.), p. 190 -224.

Weissenborn, J. (1992). "Null-Subjects in Early Grammars: Implications for Parameter Setting Theories." In: Weissenborn, J., H. Goodluck & T. Roeper (eds.)., p. 269-299.

Weist, R. M. (1983a). "Prefix versus Suffix Information Processing in the Comprehension of Tense and Aspect." *Journal of Child Language* 10 (1983), p. 85-96.

Weist, R. M. (1983b). "The Word Order Myth." *Journal of Child Language* 10 (1983), p. 97-106.

Weist, R. M. (1986). "Tense and Aspect: Temporal Systems in Child Language." In: P. Fletcher & M. Garman (eds). *Language Acquisition,* 2nd. ed. Cambridge: Cambridge University Press., p. 356-374.

Weist, R. M., Wysocka, H., Witkowska-Stadnik, K., Buczowska, E. & E. Konieczna (1984). "The Defective Tense Hypothesis: On the Emergence of Tense and Aspect in Child Polish." *Journal of Child Language* 11 (1984), p. 347-374.

Wexler, K. (1994). "Optional Infinitives, Head Movement and the Economy of Derivation in Child Grammar." In: N. Hornstein & D. Lightfoot (eds.), p. 305-350.

Wexler, K. (1995). TNS, DET and Feature Strength and Interpretability in Early Child Grammar. Talk and Handout at the *Workshop on Clausal Architecture,* Bergamo / Italy, 17.11.1995.

Wexler, K. (1998). "Very Early Parameter Setting and the Unique Checking Constraint: A New Explanation for the Optional Infinitive Stage." *Lingua* 106, p. 23-79.

Wexler, K. (1999). "Very Early Parameter Setting and the Unique Checking Constraint: A New Explanation for the Optional Infinitive Stage." In: A. Sorace, C. Heycock & R. Shillock (eds.) *Language Acquisition: Knowledge Representation and Processing.* Amsterdam: North-Holland..

Wexler, K. (2002). "Strong and Surprising Cross-Linguistic Developmental Differences in the Omission of Object Clitics Follow from the Unique Checking Constraint." Paper presented at the ESF Conference on Theoretical and Experimental Linguistics: The Syntax of Normal and Impaired Language. Corinth, Greece, June 2, 2002.

Wijnen, F. (1994). "Incremental Acquisition of Phrase Structure: A Longitudinal Analysis of Verb Placement in Dutch Child Language." Ms. Universiteit Groningen.

Wijnen, F. (1996). "Temporal Reference and Eventivity in Root Infinitives." *MIT Occasional Papers in Linguistics,* No. 12, p. 1-25.

Willim, E. (1986). *Some Aspects of Word Order in English and Polish: A Government and Binding Study.* Doctoral Dissertation, Jagiellonian University, Krakow.

Witkoś, J. (1993a). *Some Aspects of Phrasal Movement in English and Polish.* Doctoral Dissertation, Adam Mickiewicz University, Poznań.

Witkoś, J. (1993b). "The Theory of X-Binding and Some Aspects of Scrambling in Polish." In: *Papers and Studies in Contrastive Linguistics (The Polish-English Contrastive Project)*, vol. 28, Adam Mickiewicz University, Poznań.

Witkoś, J. (1998). *The Syntax of Clitics: Steps Towards a Minimalist Account.* Poznań: Motivex.

Zabrocki, T. (1981). *Lexical Rules of Semantic Interpretation. Control and NP-Movement in English and Polish.* Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.

Zabrocki, T. (1990). "Adjunction to IP and NP: Evidence from Polish." Ms., Adam Mickiewicz University, Poznań.

Zusammenfassung

Die vorliegende Dissertation mit dem Titel „Optional Structures in the Acquisition of Polish: A Cross-Linguistic Perspective" beschäftigt sich in einer komparativen Analyse mit einem momentan viel beachteten Gegenstand der Spracherwerbsforschung: dem Phänomen der Optionalität beim frühkindlichen Grammatikerwerb.

In der linguistischen Theoriebildung im Rahmen von Chomskys „generativer Grammatik" zeichnet sich in den letzten Jahren ein zunehmendes Interesse an einem interdisziplinären Unternehmen ab, das Rizzi als „in-depth collaborations of theoretical linguists and developmental psycholinguists" umschreibt. Diese Zusammenarbeit von Grammatiktheorie und Spracherwerbsforschung erscheint in zweierlei Hinsicht vielversprechend:

Zum einen stellen Ergebnisse aus Spracherwerbsuntersuchungen eine breite(re) empirische Basis für die linguistische Theoriebildung bereit. So lassen sich anhand von Erwerbsdaten konkurrierende Modelle der Universalen Grammatik (UG) wesentlich besser testen, als dieses allein auf der Grundlage von zielsprachlichen Daten der Fall wäre.

Auf der anderen Seite können theoretische Ansätze aus der Syntaxtheorie von explanatorischer Bedeutung für zunächst undurchsichtige Eigenschaften von Spracherwerbsdaten sein. So wurde z.B. für die frühe Null-Subjekt-Phase eine Erklärung vorgeschlagen, die sich auf die Parametertheorie stützt.

Die vorliegende Arbeit versucht, dieses wechselseitige Verhältnis von linguistischer Theoriebildung und Spracherwerbsforschung zu umspannen und in einem kleinen Teilgebiet des Erstspracherwerbs – der Analyse optionaler Strukturen beim Erwerb des Polnischen – konkret zu realisieren.

Die Arbeit bewegt sich im theoretischen Rahmen des *Prinzipien- und Parameter-Modells,* demzufolge die menschliche Sprachfähigkeit – als Universalgrammatik (UG) bezeichnet – zum einen aus abstrakten, unveränderlichen *Prinzipien* besteht, und zum anderen aus variablen *Parametern,* die einzelsprachlich unterschiedlich besetzt sind, wodurch strukturelle

Unterschiede zwischen den Sprachen erklärt werden können. Dieser Ansatz wird in Kapitel 2 diskutiert.

Auf der Suche nach einer sprachübergreifenden Darstellung und Erklärung der menschlichen Sprachfähigkeit liefern neuere Untersuchungen im Bereich des Erstspracherwerbs wichtige Evidenz für die Frage, welche Rolle der Universalgrammatik in der frühkindlichen Grammatik zukommt und inwieweit Kinder schon in den frühen Phasen der Sprachentwicklung über sprachliche Strukturen verfügen, die dem zielsprachlichen System nahekommen und somit für eine *Kontinuitätshypothese* sprechen.

Die Frage nach der Architektur der frühkindlichen Grammatik (insbesondere der Präsenz funktionaler Kategorien) lässt sich nur aufgrund von empirischer Evidenz in Form von Erstspracherwerbsdaten beantworten. Untersuchungen von Kindersprache haben seit ein paar Jahren das Phänomen der Optionalität in der frühen Kindersprache beobachtet. Ein viel zitiertes Beispiel hierfür ist der Bereich der Verbalflexion. In verschiedenen germanischen Sprachen (z.B. Deutsch, Englisch, Niederländisch, den Festland-skandinavischen Sprachen) treten in frühen Entwicklungsstufen (im Alter von zwei Jahren etwa) sowohl finite als auch nicht-finite Verbformen in Hauptsätzen auf, obwohl diese Strukturen in der Erwachsenen-Grammatik nicht zulässig sind. Diese Phase der Oszillation zwischen finiten und nicht-finiten Formen zu einem Zeitpunkt, an dem bereits syntaktische, abstrakte Prinzipien wie z.B. Kopfbewegung, nachweisbar sind, wurde seit Wexler (1994) mit dem Begriff *Optional Infinitive Stage* belegt (s. die Diskussion in Kapitel 3).

Es stellt sich nun die Frage, inwieweit optionale Strukturen allgemein – und hier konkret die *Optional Infinitive*-Phase – ein generelles Phänomen der frühkindlichen Grammatik darstellen. Hier kommt sprachvergleichenden Untersuchungen eine entscheidende Bedeutung zu. Im Unterschied zu den Befunden für die germanischen Sprachen haben Untersuchungen zu anderen Sprachen (wie etwa Italienisch, Spanisch oder Türkisch) gezeigt, dass die untersuchten Korpora aus diesen Sprachen entweder keine oder aber nur eine sehr begrenzte *Optional Infinitive*-Phase aufweisen.

In der vorliegenden Arbeit wurde nun die Gruppe der slawischen Sprachen (exemplarisch anhand des Polnischen, und in Ergänzung hierzu des Russischen und Kroatischen) analysiert.

Die slawische Sprachengruppe bildet insbesondere wegen ihrer differenzierten Flexionsparadigmata einen interessanten Untersuchungsgegenstand, gerade im Hinblick auf die zu untersuchende Verbmorphologie (s. Kapitel 4).

Die empirische Grundlage dieser Dissertation bildet eine Langzeitstudie, die den Spracherwerbsverlauf von drei polnischen Kindern im Alter von 1;11 bis 3;3 dokumentiert. Ich habe dazu in Gdansk/Polen spontansprachliche Kindersprachdaten erhoben, die eine aus drei Korpora bestehende Datenbasis (*Dagmara-*, *Anna-* und *Aleksandra*-Korpus) mit mehr als 11.500 analysierbaren Äußerungen ergeben hat. Um sicherzustellen, dass die aufgenommenen und analysierten Daten den Kriterien der Spontansprache entsprachen, entwickelte ich zunächst ein standardisiertes Verfahren für die Aufnahmen sowie ein Schema für die Transliteration, die Kodierung sowie die anschließende Auswertung (s. Kapitel 5 der Arbeit). Von besonderem Interesse ist der *Aleksandra*-Korpus, der einen langen Zeitraum von mehr als zwei Jahren lückenlos und kontinuierlich abdeckt, und sogar die frühesten Phasen der sprachlichen Entwicklung (Ein- und Zwei-Wort-Phase, ab 1;4 Jahren) erfasst.

Die von mir erhobenen polnischen Kindersprachdaten wurden dann im Rahmen einer sprachvergleichenden Diskussion auf die folgenden drei Forschungsfragen hin untersucht:

1. Lässt sich in den polnischen Daten ebenfalls eine Phase der optionalen Infinitive nachweisen, und wie sieht die Verteilung von Finitheit aus?
2. Welche Rückschlüsse erlauben die frühesten Phasen der Sprachentwicklung (d.h. die Ein- und Zwei-Wort-Phase des *Aleksandra*-Korpus)?
3. Wie sind Subjekte realisiert und wie sieht die Verteilung overte Subjekte vs. Null-Subjekte aus?

Darüber hinaus wurden alle drei Fragen jeweils aus einem komparativen Blickwinkel heraus analysiert und im Lichte des Phänomens der Optionalität generell diskutiert.

Die quantitative und qualitative Untersuchung der Daten hat zu folgenden Ergebnissen geführt, die in Kapitel 6 ausführlich dargestellt sind:

Die polnischen Spracherwerbsdaten zeigen keine Evidenz für eine *Optional Infinitive*-Phase im Polnischen. Die (wenigen) auftretenden Hauptsatz-Infinitive wurden in einer Kontext-

Analyse untersucht, und weisen vorwiegend modale Bedeutungen oder elliptische Funktionen auf, sind also kontext-lizensiert.

Auch in den frühsten Entwicklungsstufen zeigt sich kein Anhaltspunkt für das Auftreten einer *Optional Infinitive*-Phase, wohl aber ein anderes unerwartetes Phänomen: das Auftreten von Imperativen in Deklarativ-Kontexten („*Declarative Imperatives")* während der Ein- und Zwei-Wort-Phase, bis zum Alter von 1;8. Da dieser abweichende Gebrauch von Imperativen mit der gleichzeitigen Beherrschung grammatischer Prozesse wie Kongruenz und Finitheitsdistinktion einhergeht, habe ich diese Strukturen als morphologische Ersatzformen analysiert, die dennoch temporal interpretierbar sind. Unmittelbare Unterstützung für diese These liefern Beobachtungen aus anderen Sprachen, wie z.B, Partizipien im Griechischen sowie in Inuktitut, sowie ähnliche Default-/Ersatzformen im Kroatischen, Russischen und Österreichischen Deutsch.

Was die Distribution von overten vs. nicht-realisierten (=null-)Subjekten anbetrifft, so zeigen die polnischen Daten einen hohen Anteil an overten Subjekts-Strukturen, was angesichts des Null-Subjekt-Status' des Polnischen überraschend ist. Darüber hinaus zeigt eine weitere Differenzierung der overten Subjekte, die ich noch auf ihren Subjekt-Typ (nominal vs. pronominal) untersucht habe, einen ausgesprochen hohen Anteil an pronominalen Subjekten (80% im Durchschnitt) gegenüber lexikalischen Subjekten (NPs). Diese Subjekt-Pronomina scheinen insbesondere mit der Verwendung der Kopula *sein* zu korrelieren. Auch für dieses Ergebnis gibt es parallele Ergebnisse für das Kroatische.

Die Ergebnisse meiner Untersuchung sowie ihre komparative Deutung (s. Kapitel 7) belegen eine sprachübergreifende Evidenz für die Präsenz funktionaler Kategorien bereits in den frühen Stadien des kindlichen Grammatiksystems, und somit für die eingangs dargestellte Kontinuitätshypothese.

Abstract

This thesis with the title “Optional Structures in the Acquisition of Polish: A Cross-Linguistic Perspective” focuses on the phenomenon of optionality in early child grammar.
The last two decades have seen a development in the theory of first language acquisition that Rizzi describes as “in-depth collaborations of theoretical linguists and developmental psycholinguists”. The growing interest in language development coincided with the consolidation of the *Principles and Parameters* model of Universal Grammar (UG) in the framework of Chomsky’s theory of generative grammar in the mid eighties.

This cooperative venture of grammatical theory and language acquisition studies seems to be promising in two respects: On the one hand, results of language acquisition studies provide a broader empirical basis, thereby enabling the researcher to test competing linguistic theories.
On the other hand, parametric models introduced a theoretical framework well adapted for the comparison of systems basically cast in the same mould, but diverging on some structurally well-defined points. The same methodology used for comparative studies of adult languages can be extended to child language.

This thesis attempts to capture the cooperation of linguistic theory and language acquisition studies in an analysis of optional structures in the acquisition of Polish. The analysis is couched in the *Principles and Parameters* framework of language acquisition studies, according to which the human language faculty – referred to as *Universal Grammar* (UG) – consists of two parts: on the one hand, there are abstract, invariable principles which are supposed to be innate. What is open to variation are the parametric values that the principles can take on. The Principles and Parameters conception of UG allows for cross-linguistic variation by associating with the principles of UG a small number of parameters of variation.

Any approach to the nature of the language faculty involves an account of the shape of the linguistic knowledge in the brain and also an explanation of how this knowledge enters the brain. Language acquisition studies are expected to provide evidence in order to answer the question of the early representation of grammatical knowledge in the child’s grammar. A *Continuity View* of language development amounts to positing that children’s grammars include the same functional projections as adults’ grammars, and that the early clausal architecture of the child’s grammatical system corresponds to the adult one.

This view is challenged by the phenomenon of *optionality* in early child grammar. A major case of optionality is found in the area of verbal inflection: there is a stage (around the age of two) in which two verbal forms seem to coexist in declarative main clauses: the adult-like finite form and an optional (root) infinitive. In various Germanic languages (e.g. German, English, Dutch and Mainland-Scandinavian languages) children at an early age produce infinitival constructions as main clauses, which is ungrammatical in the target languages. Following Wexler (1994), this oscillation between finite and non-finite forms, which can be shown to happen at a stage where abstract syntactic principles are already present, has been named *Optional Infinitive Stage* (see discussion in chapter 3).

On the other hand, there are languages (such as Italian, Spanish or Turkish) which show no or only a very limited *Optional Infinitive Stage.* Therefore, the question arises in how far optional structures represent a general phenomenon of early child grammar.

The aim of this thesis is to contribute to the cross-linguistic discussion on Ois by providing empirical evidence from Polish, a rich-agreement language, which has not yet been analysed in this respect. Essentially, each form of the verbal paradigm in Polish is distinctly marked for person, number, and, in some tenses, also gender. In contrast to the Germanic languages, the infinitive represents a highly marked form in the Polish paradigm. Therefore, the study of the OI phenomenon in Polish seems to be of both empirical and theoretical interest: in addition to extending the previous empirical findings to a Slavic language, it might shed light on the theoretical cross-linguistic OI-generalization proposed in the literature.

This thesis is based on a longitudinal study of Polish child language which I collected in Gdansk / Poland over a period of three years. The aim of this empirical study was to collect naturalistic, conversational data (matching the *spontaneous speech* criteria) from three children. The data base I obtained consists of three sets of data, the *Dagmara-, Anna-,* and *Aleksandra*-Corpus with a total of more than 11,500 analysable utterances. In order to make sure that the data would meet the spontaneous speech criteria, I developed a standardized procedure of data recording, transliteration, encoding and data classification (see description of my *data evaluation scheme* chapter 5).

Of the three corpora, the *Aleksandra*-Corpus is of special interest since it covers data from the very beginning of child language production (one- and two-word-stage, beginning at the age of 1;4), thereby providing an insight into the earliest stages of language development.

I analysed the data according to the following research questions:

1. Do the Polish child language data show any evidence for an *Optional Infinitive Stage?* What does the distribution of finiteness look like in the three corpora?
2. What do the earliest stages of acquisition (i.e. one- and two word stage, as represented in the *Aleksandra*-Corpus) reveal about the clausal architecture of early child grammar?
3. How does the realization of subjects develop and what does the distribution of null- vs. Overt subjects look like?

In addition to this, all three questions were analysed from a cross-linguistic point of view in order to shed light on the theoretical concept of *optionality* in children's early grammatical systems.

The quantitative and qualitative analysis of the data led to the following results which are presented in detail in chapter 6:

The Polish child data show no evidence for an *Optional Infinitive Stage* in Polish. The few occurring main clause infinitives were analysed in a context analysis which showed that main clause infinitives mainly occur (if at all) in modal, volitional or elliptical contexts, which amounts to saying that they are contextually licensed.

Instead, an investigation of the earliest stages of acquisition, the one- and two-word stage, approximately until the age of 1;8, has revealed an interesting, unexpected finding: the use of imperatives in declarative contexts , associated with a descriptive meaning ("*Descriptive Imperatives*"). At the same time, the Polish children analysed here seem to have a good command of subject-verb agreement and other items of morphosyntactic knowledge. Having considered some cross-linguistic evidence for a similar kind of "default" use of *Descriptive Imperatives* in Russian, Croatian and Austrian German child language, as well as participles

in Greek or Inuktitut, I analysed these structures as *surrogate verb forms* that can still have a temporal interpretation.

As far as the distribution of overt- vs. null subjects is concerned, the Polish data show a high frequency of overt subjects, which is surprising given the status of Polish as a null subject language. Therefore, it seemed worthwhile to investigate the *type* of subject involved (lexical, i.e. N/NP, vs. pronominal subjects). The distributional analysis revealed that the rate of nominal/lexical subjects is strikingly low compared to the abundance of pronominal subjects (80% on average, notably personal pronouns). These subject pronouns tend to correlate with the use of copular constructions. This finding has been confirmed by similar observations for Croatian child language.

All phenomena analysed so far as well as their interpretation from a cross-linguistic perspective seem to speak in favour of differentiated morphosyntactic knowledge and the presence of functional categories in the early child grammar, thereby supporting a *Continuity View* on language development.