SOAS Working Papers in Linguistics Vol. 12 (2002): 315-333

Aspects of the acquisition of functional projections in embedded clauses[1]

Thomas Doukas
129669@soas.ac.uk

Abstract

The purpose of this paper is to give a brief account of the acquisition of embedded clauses in children's language. For the purpose of this study corpora of naturalistic speech of children speaking Modern Greek, Italian and English have examined. My aim is to, first give evidence for the fact that embedded clauses are acquired cross-linguistically earlier than generally assumed, and second provide support for the Maturational Hypothesis of Language Acquisition. Statistical and numerical results are given in this work concerning the use and proportion of such embedded clauses. I will show that the three children examined for the purposes of this paper use embedded clauses that contain functional categories, following the adult model of grammar. Furthermore, I will argue and show through the analysis that infinitives and that-clauses are consistently used more than other embedding material due to the fact that they are verbal arguments.

0 Introduction

Diverse studies on language acquisition in many languages have demonstrated that, children between the ages of 2 and 3 years have already acquired different aspects of grammar, and they appreciate the syntactic value of such grammar. Nevertheless, it has been claimed that children's language and grammar is not identical with that of adults and that the former need to restructure their grammar in order to adopt the adult model.

The main goal of the present study is to analyse some aspects of the acquisition and status of embedded clauses in Modern Greek, Italian and English placing such results within the theories stated above. Being more specific, the two main tasks of this research are: First, to give evidence and account for the fact that embedded clauses are acquired cross-linguistically earlier, than generally assumed, and second to provide support for the Maturational Hypothesis, and against the Discontinuity Hypothesis of Language Acquisition.

The first section of this work offers a brief presentation of prominent theories concerning the acquisition of embedded clauses. More specifically I will focus on the debate concerning the status of embedded clause in the literature (Discontinuity Hypothesis and Maturational Hypothesis).

The second section contains a brief introduction of primary naturalistic speech data collated by myself (of a Modern Greek-speaking child, called Maria, aged between 2;0,24 (i.e. 2 years, 0 months and 24 days of age) and 2;8,27), of the way the data have been collated and the way these have been used for the present analysis. This therefore entails the presentation of the corpus of the naturalistic speech, the methodology and the processes used for the classification of the utilized utterances, and finally, the extraction of specific parts of the data. Furthermore, in the section I

[1] I would like to express my gratitude to Andrew Simpson for all the helpful comments and support during the elaboration of this paper. I am also grateful to Dina Chaidou and Vina Tsakali for many fruitful discussions.

will present the rest of the data used for this analysis, names, sources, classification, and a typological key for the different kind of embedded clauses used for this work will be presented too.

In the third section I will analyse the speech production of all the examined data, mentioned above. In this part, numerical and statistical presentation of the data and a first approach and placing of the findings within the theory and the different studies on language acquisition will be introduced. Furthermore, the section provides an analysis and first account of the early grammar's constituents, placed within the supported theory, i.e. the maturation hypothesis.

Finally, the last part of this work contains the conclusions of the analysis, the comparison between the status of embedded clauses within the examined data, and the general conclusion within the general framework and theory of embedded clauses from the literature. In this part I will propose that the emergence and first use of embedded clauses and complex constructions comes earlier in children's language acquisition process than is generally proposed.

1 The Discontinuity Hypothesis

As stated in the introduction above, children's grammar differs from adults' with respect to diverse aspects. The passage from the initial stage (S0) or Early Grammar (G1) of infants' grammar to the Stable Stage (Ss) of adults' grammar involves a process of restructuring.

In this subsection I will discuss and give the main points of view of a hypothesis of acquisitional development, different and in part critical of the model of Maturation adopted in this work. Such a hypothesis has been called the discontinuity, or non-instantaneous hypothesis (term as used by Hyams, 1986), or structure-changing hypothesis (term used by Guasti, 2002). These hypotheses differ minimally from each other, and all claim that early grammars are subject to strong and radical restructuring of their principal constituents during attainment of the adult grammar.

The discontinuous model of development, in particular, the so-called "semantically-based child grammars", entails the existence of a rather radical restructuring from a semantically-based child grammar to a syntactically-based adult grammar[2].

According to this theory of semantically-based child grammars the early grammars map underlying semantic categories, for example agent, action, entity, attribute, etc., directly onto the linear position in a surface expression. These grammars do not contain the syntactic categories, relations, or rules, which define the adult system. Hierarchical structure is also assumed to be absent. The semantic categories are assumed to be universal, innately available to the child by virtue of his general conceptual system (Hyams, 1986).

As a consequence, a semantically-based model of early grammar lacks a syntactic system. Furthermore, it has been proposed that the semantically-based systems persist until age of three or four.

The prediction for embedded clauses therefore, according to the above outlined theory, is that children's grammar is completely different from that of adults in the sense that it does not contain any syntactic and functional structure and that the system is subject to a significant restructuring.

[2] I will show, however, that such semantically-based system is empirically inadequate as a theory of early linguistic competence, and hence that it does not constitute evidence for discontinuous development.

1.1 The Structure-Change Hypothesis and Embeddedness

In studying the development of embedded clauses, the notion of Control is particularly relevant because it accounts for the referential properties of the subject of a clause. Control refers to the phenomenon in which the understood (PRO) subject of an infinitival clause must be anaphorically dependent on a specific argument of the matrix clause. In Haegeman (1994):

> The term **control** is used to refer to a relation of referential dependency between and unexpressed subject (the **controlled** element) and an expressed or unexpressed constituent (the **controller**). Those of the controller ...determine the referential properties of the controlled element. (Haegeman, 1994, p. 263)

Control Theory is based upon the notion of the Extended Projection Principle, where the subject positions must be syntactically represented; in other words all projections of IP have a subject who must be projected. Moreover, Control Theory explains those aspects of the behaviour of PRO which are not captured by other theories (such as Binding Theory). Thus, the distribution and interpretation of PRO, especially for infinitive clauses is very important for the development of embedded clauses and is regulated by the module of the grammar called Control Theory.

Relevant to the nature of control, another hypothesis of the discontinuous acquisition model, discussed in Guasti (2002) is the structure-changing hypothesis, which is contrary to the maturational model of development. Within this hypothesis it is argued that children move through different grammars of control before they achieve adult-like competence, and hence children make mistakes in interpreting syntactic structures, because they mentally represent sentences incorrectly. In particular, concerning Control Theory and principles of embedded sentences, Guasti (2002) claims:

> Children manifest developmental changes in control because in approaching the adult target, they have to learn the lexical properties of verbs and of (subordinating) conjunctions; lacking full knowledge of these properties, they make different hypotheses about the attachment site of subordinate clauses. (Guasti, 2002)

The proposed model suggests that children go through four different stages of control development and of interpretation of PRO. Thus, in Grammar 1, PRO is allowed to be interpreted freely in all complement and adjunct constructions. In Grammar 2, free interpretation of PRO is allowed only for adjunct clauses, while in Grammar 3 and 4 starting with object control of PRO and then with subject control of PRO, children's grammar comes closer to the target adult grammar.

It is claimed for this early grammar, namely 1 and 2, that children do not have access to a recursive rule for embedding clauses and thus they analyze control structures as coordinate structures. After Grammar 2, children's development incorporates the recursive rule for embedding clauses, starting with complements first and then in Grammar 3 and 4 with adjunct constructions as well, abandoning completely the coordinate structures used in the beginning.

What is important to mention for the analysis of this work, noticed in Guasti (2002) is that children do not necessarily go through each of these grammars successively but they may skip one of the grammars.

The last of the hypotheses operates according to a similar line of reasoning, namely the non-instantaneous hypothesis (Hyams, 1986), where it is proposed that children have no immediate access to all the principles of UG and all the primary linguistic data.

1.2 The Maturation Hypothesis

The maturational hypothesis proposed by Borer & Wexler (1987) is closely related to the continuous, instantaneous and other models. Borer & Wexler (1987) argue that grammatical development is a continuous process in that it is constrained by principles and parameters of UG. This development is subject to maturation, in the sense that not all principles of UG are specified at the initial state. Rather, it is claimed that early grammars are constrained by those principles specified.

Consequently certain aspects of grammatical development may be delayed because of the inability to analyze particular data, or the maturation of specific principles of grammar. The maturation of various processing related abilities such as memory or attention, as well as development in the conceptual domain, may also be responsible for delays.

Additionally, Wexler (1992) suggests that children's grammar for control undergoes developmental changes and that non-adult responses include just two stages (contrary to the four stages proposed in the structure-changing hypothesis). According to Wexler (1992), in early children's linguistic development PRO is not accessible to them; it becomes available upon maturation around 3-4 years.

During Stage 1 children do not have access to PRO and thus allow free interpretation of PRO in non-finite complement and adjunct clauses. In this Stage children know how to build complex structures and know the recursive rule for embedding complements and adjuncts. They also know that every clause must have a subject. However, since PRO is not available to them, children must reanalyse subordinate non-finite clauses and non-finite adjuncts, in a way that avoids the use of PRO and hence as NPs that do not need the presence of PRO. In fact NPs, unlike clauses, do not require a structural subject. In other words, children do not have access to PRO and thus they are forced to analyse non-finite complements and adjunct as NPs.

On the other hand, during Stage 2, children have access to PRO. They interpret PRO as adults do when it occurs in non-finite complement clauses. However, they still allow free interpretation of PRO in non-finite adjunct clauses. Thus, they assign a clausal representation to the complement of verbs, fill the subject position of this non-finite clause with PRO, and interpret it correctly.

As a consequence of the fact that the early grammar differs from the adult grammar, the former must be altered or restructured during the acquisition process. Within the framework adopted in this work, this restructuring involves a resetting of the parameters, which is done on the basis of positive evidence in the child's linguistic environment.

A central hypothesis in the framework is that grammatical development is continuous, in the sense that while grammars do undergo restructuring during the course of development, the restructuring is within narrowly defined limits – those imposed by the principles and parameters of UG (Poletto, 2000). Thus, as stated in Hyams (1986) early grammar is delearnable.

Guasti, (2002), Hyams, (1986 & 1985), Lightfoot (1991), Crain & Thornton (1998), Crain & Lillo-Martin (1999), Ingram (1989) among others, propose that this

distinction of early grammar from the adult one is due to the phenomenon of learning of particular lexical items and their associated properties, for example, meaning, subcategorization, grammatical category, etc. rather than due to a lack of UG items. Lexical items must be learned largely on an "item by item" basis. Thus, once again the child may need greater exposure to data to learn those properties, which are idiosyncratic to particular lexical items. For example, there are verbs, which select propositions as arguments, e.g. think, tell, want, etc. These verbs also select particular complementizers, e.g. that, for. It seems clear that the child will not control complex sentences until it learns these verbs and their selectional requirements (in Guasti, 2002).

Hyams, (1986 & 1985) claims, with regard to this last property proposed, that the early grammar exerts a filtering effect on the data. The child systematically ignores a class of elements for which his grammar provides no structural description.

> This particular paradox is an instance of the much more general phenomenon of "selective attention". It is well known that children do not analyze all of the available data (if they did acquisition would indeed be instantaneous as in the familiar idealization). Rather they select certain data for analysis and ignore other... it is well known that early language centers largely on objects and events in the immediate environment. (Hyams, 1986, p94, 97)

1.3 Embedded clausses and the Maturation Hypothesis

As stated above, the embedding principle is subject to the maturational process as well. In fact Wexler (1992) proposes two stages for the theory of control, where children's grammar during the second stage develops the interpretation and analysis of PRO as in the adult-like model. Thus, the embedding principle is subject to maturation and restructuring.

Hyams (1986 & 1985) proposes that all restructuring is induced by the acquisition of particular lexical items (and their associated properties). For example, it may be the case that embedding is introduced into the grammar by the acquisition of verbs which subcategorize a sentential complement (Hyams, 1985). Furthermore, based on a study of Limber (1973), she notes that complements with a given complement-taking verb will appear within a month after the first use of that verb in any construction:

> He notes that the first complex sentences contain verbs such as *want* and *watch* which the child previously used with NP objects. Thus, the acquisition of particular verbs induces a restructuring in the grammar to include sentential embedding under VP. (Hyams, 1986 p93)

In this respect their grammar for control is like the adult grammar from the start. There is no point during development in which children lack PRO or ignore the rules governing the interpretation of PRO. It is not knowledge of control that develops, but knowledge of lexical properties of verbs and conjunctions and knowledge of the way clauses are arranged.

> The lexicon is acquired, in large part, on an item-to-item basis, and it seems reasonable to assume that the child's knowledge of form-class

> relationships is one of the factors, which aid him in determining the grammatical category of particular lexical items. (Hyams, 1986 p79)

We noted previously that the idealization of instantaneous acquisition involves the assumption that the child has immediate access to all the principles of UG and all the primary linguistic data. In that case, it is obvious that it is the maturation of the child's representational abilities, which enable him to consider data, which he previously, ignored (Hyams, 1986 & 1985). Thus, the non-adult interpretation of control complements and adjuncts is due to failure to properly represent these constructions rather than to defective knowledge of control principles.

2 Data: *Maria's* Corpus (Modern Greek)

Part of the analysed sentences in this work has been collated from primary data of personal labour. This piece of data contains the speech production of a child of Modern Greek language, called Maria. The data was collated in the following way: Maria was tape-recorded starting from the age of 2;0.24 (i.e. 2 years, 0 months and 24 days of age) until the age of 2;8.27 for a total of eight tape-recordings, with frequency almost one recording per month of variable duration between 30 and 45 minutes.

The recordings were realised in my presence, where it was possible, if not, in the presence of well instructed relatives. The tapes, containing the speech production of Maria were made in her own home, in an environment familiar to her.

All the recordings of the tapes have been transcribed by myself, partly following the transcription model called CHAT, used for the transcriptions of data of children contained in the data base of CHILDES (Mac Whinney & Snow, 1985).

I have used an appropriate adopted alphabet[3] that mainly corresponds to other languages in order to transcribe the sentences in a clear way.

As the recordings were audio-tapes, I had to deal with diverse problems during the transcription, e.g. understanding things and objects that Maria took or showed. For this reason in the transcriptions, there are a lot of cases of incomprehensibility, some resolved with the help and guidance of the relatives present in the recordings.

Finally the recordings were made under circumstances of playing or reading fairy tales. General information concerning each recording are presented in the following table:

Maria	Age	Total sentences	Sentences used
1	2;0,24	84	47
2	2;2,8	149	87
3	2;3,18	188	138
4	2;5,4	183	136
5	2;5,24	240	122
6	2;7,1	139	80
7	2;8,3	118	66
8	2;8,27	160	106
Total		**1261**	**782 (62%)**

Table1: Corpus of Maria

[3] The proposed alphabet is created in order to provided easy reading of the examples and corpus, and it is based mainly in phonological transcription of words which do not correspond to Modern Greek orthography.

The table contains the distribution of the sentences used in the research, out of the total of sentences of the whole corpus. It was only possible to utilise 62% of the sentences due to incomprehensibility of the rest of sentences.

2.1 Other children Corpora (English and Italian)

Apart from Maria's corpus, two other corpora of naturalistic speech from the CHILDES database (Mac Whinney & Snow, 1985) have been analysed for the purpose of the present study. The information from the other corpora used is as shown in the following list:

- For Italian language data:

Tonelli.zip→ Sara→ 2;0.20 – 2;8.21 12 files

- For English language data:

Clark.zip→ Shem→2;2.16 –2;8.29 23 files

The main criterion used for the selection[4] of the different children is age, and hence children whose age matches that of Maria have been selected.

2.2 Production part methods and typological facts

The extraction of the data and all the analysis were realised manually and unclear sequences or those missing relevant parts of speech i.e. those containing xxx[5] in the transcriptions were omitted.

All songs occurring in the data were ignored, because they were considered to be memorised. On the other hand all imitations, total or partial, of phrases produced by the adults were considered.

For this analysis, I considered embedded clauses of different types, according to the use of such clauses by the children in their speech production. The typology of such embedded clauses is presented in the following table, divided by language and type of sentence. The first set of rows, under the title "arguments of the verbs", contains infinitives and that-clauses, and the introductory particles of such constructions. Those two types are the more attested in the analysed data due to the fact that they are verbal arguments. The following rows contain verbal adjuncts embedded construction, i.e. relative clause (RCs), conditionals, and purpose. Again the introductory particles for each type are given. RCs are highly attested too. On the other hand, conditional and purpose sentences are rarely attested in the corpora of all the children, but for the sake of their presence in some of the data, they were counted in the typology. Thus, the most attested forms of embedded clauses in the analysed data, are the first three types, i.e. infinitives, that-clauses and RCs.

[4] The selection of those children form the CHILDES was based upon the following reasons:
Some of the children of the data do not match Maria's age at all. In other zips, the age of the recorded child has not been specified, and thus it was impossible to say exactly what the age is. Several children are analysed either for phonological or morphological analysis. It is very confusing to read and analyze these files because the conversation is always interrupted by phonological and morphological comments and the sequence is not clear. Diverse zip data contain speech production of groups of children playing together, usually children of more advanced age.

[5] The symbol xxx indicates incomprehensible or unintelligible speech, not treated as a word.

	TYPE/LANGUAGE	ENGLISH	GREEK	ITALIAN
EMBEDDED CLAUSES/ ARGUMENTS OF THE VERB	INFINITIVES	to	na+subjunctive[6]	v+(preposition)+infinitive
	THAT-CLAUSES	that	oti+indicative	che+indicative/subjunctive
EMBEDDED CLAUSES/ ADJUNCTS	RELATIVE CLAUSES	that/(0)	ton opoio/(pu)	il quale/che
		who/whom/that	ton opoio/(pu)	di chi/il quale/che
		which/that	ton opoio/(pu)	il quale/che
		where/that	pu	dove
	CONDITIONAL	if	an/ama	se
	PURPOSE	in order to	ja na	per/a/da+infinitive

Table 2: Typology of embedded clauses in the languages used for this work

3 Corpora analyses (Maria, Sara, Shem)

In what follows there is a brief numerical and statistical presentation together with various details of the data used for the analysis.

From the transcribed registrations[7], I have isolated all embedded clauses from the diverse corpora I used for the comparison. The following tables concern the distribution of embedded clauses of the three children.

Child/File	Age	Sentences used	Embedded clauses
Maria1	2;0,24	47	2 (4%)
Maria2	2;2,8	87	14 (16%)
Maria3	2;3,18	138	8 (6%)
Maria4	2;5,4	136	19 (14%)
Maria5	2;5,24	122	5 (4%)
Maria6	2;7,1	80	8 (10%)
Maria7	2;8,3	66	10 (15%)
Maria8	2;8,27	106	9 (8%)
Total		**782**	**75 (10%)**

Table 3: Occurrences and percentages of embedded clauses out of the total of used sentences in Maria's corpus

[6] The Modern Greek language has no infinitives, but the form presented in the table of the particle "na" plus subjunctive mood. In various studies, it has been claimed that this form corresponds to the infinitives of other languages, but the idea has been abandoned. Nevertheless, I consider this form to be the only correspondance of infinitives even if it is not a proper infinitive.

[7] It is important to mention that the quantity of sentences reported in the tables for each file heavily depends on the quantity and length of each file. In other words longer files contains more sentences and hence proper figures are not indicative for progress. For this reason the percentage of the occurances of embedded clauses is given next to each figure.

The same distribution of embedded clauses is given for Sara and Shem[8] respectively in the following tables.

Child/File	Age	Sentences used	Embedded clauses
Sara1	2;0.20	113	2 (2%)
Sara2	2;1.15	125	10 (8%)
Sara3	2;2.1	142	17 (12%)
Sara4[9]	2;2.11	44	3 (7%)
Sara5	2;3.2	141	8 (6%)
Sara6	2;3.28	160	7 (4.%)
Sara7	2;4.18	172	12 (7%)
Sara8	2;5.9	157	12 (8%)
Sara9	2;6.13	163	17 (10%)
Sara10	2;7.10	183	27 (15%)
Sara11	2;8	165	26 (16%)
Sara12	2;8.21	118	18 (15%)
Total		**1683**	**159 (9,5%)**

Table 4: Occurrences and percentages of embedded clauses out of the total of used sentences in Sara's corpus

Child/File	Age	Sentences used	Embedded clauses
Shem1	2;2.16	198	16 (8%)
Shem2	2;2.23	193	16 (8%)
Shem3	2;3.2	240	22 (9%)
Shem4	2;3.16	170	16 (9%)
Shem5	2;3.21	182	22 (12%)
Shem6	2;3.28	213	9 (4%)
Shem7	2;4.4	199	14 (7%)
Shem8	2;4.20	184	9 (5%)
Shem9	2;4.25	-	-
Shem10	2;5.2	140	12 (9%)
Shem11	2;5.9	-	-
Shem12	2;5.16	272	35 (13%)
Shem13	2;5.23	298	32 (11%)
Shem14	2;5.30	-	-
Shem15	2;6.6	179	25 (14%)
Shem16	2;6.27	239	36 (15%)
Shem17	2;7.10	253	25 (10%)
Shem18	2;7.18	262	34 (13%)
Shem19	2;7.26	256	34 (13%)
Shem20	2;8.3	-	-
Shem21	2;8.15	260	14 (5%)
Shem22	2;8.20	279	24 (9%)
Shem23	2;8.29	173	12 (7%)
Total		**4190**	**407 (10%)**

Table 5: Occurrences and percentages of embedded clauses out of the total of used sentences in Shem's corpus

[8] During the extraction process for the data of Shem, I had to omit some of the files, because the frequency of the registrations was short and the days between two registrations were too close and thus irrelevant for the study of his speech progress.

[9] Sara's fourth registration is quite brief in comparison to the others, because it was realised by the mother of the child during holidays. In the data there is the warning that such file is of poor quality.

In table 3, 4 and 5 the instances of embedded clauses are given. The results have been reported out of the total of utterances of the three children's corpora. We can see that from a total of 782 sentences produced during eight registrations, 10% concern embedded clauses. Sara from a total of 1683 utterances produced during 12 registrations 9,5% concern the use of embedded clause while Shem the 10% out of a total of 4190 used sentences from 23 registrations. What is important to underline is that the percentages of the three children are almost the same, and hence 10% for Maria, 9.5% for Sara and 10% for Shem, a fact that suggests the early emergence and extensive use of complex constructions in their speech production.

Furthermore, all instances in all the three children reported in the tables, involve correct use of embedded contexts following the adult model. This fact suggests that children's grammar involving embedded domains is very similar if not same as adults constructions. So far the empirical evidence given in this paper supports the theory of maturation and the claim that children's constructions differ minimally from the adult model. Some examples of embedded clauses produced by the three children follow:

(1) a. pao na foreso to bufan
I-go (na) wear the jacket
I am going to wear the jacket

b. oti ime kalo pedi, lei
that I-am good child, says
Says that I am a good child

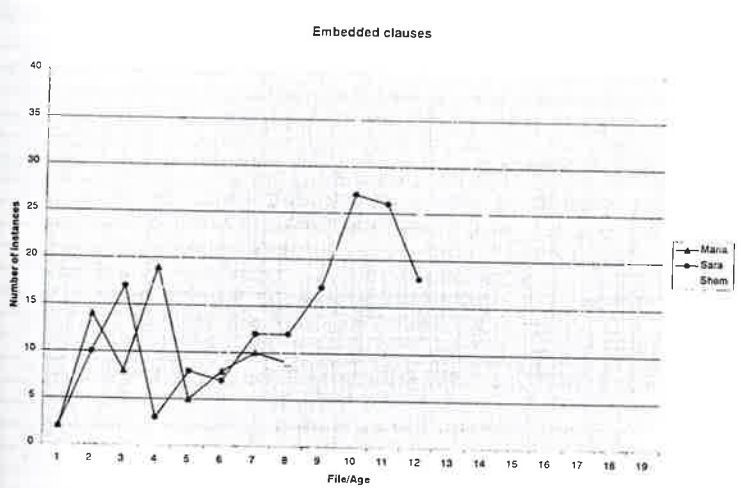
(2) a. vado a cecare [:cercare] l' alta bambola.
I-go to look for the other doll/puppet
I am going to look for the other doll/puppet

b. sai che io mi sono b(r)ucciata il ginocchio.
You-know that I (me) have burnt the knee
Do you know that I have burnt the/my knee

(3) a. I wanna show you a game.
b. I think I wanna go bed.

Examples in (1) is from Maria's corpus, while in (2) are from Sara and (3) from Shem respectively. All a. parts involve use of infinitives whereas b. cases of that-clauses. Further to the evidence presented above, based upon the numerical and statistical results of embedded clause, the examples in (1), (2) and (3), support the claim that children's complex constructions are similar to adult language. Embedded clauses produced by children at that age, and more specifically verbal argument constructions, have the same structure and contain the same functional projections with the adult language.

Finally in chart 1, above, we can see the distribution of the total embedded clauses in Maria, Sara and Shem.



In the following tables is the distribution of the different types of embedded clauses of three children's data analysed in this work. The tables were made following the typological distribution of embedded clauses given in the table 2.

Child/File	Age	Na-clauses (subjunctive)	That-clauses	RCs	Conditional	Purpose	Total
Maria1	2;0,24	1				1	2
Maria2	2;2,8	12			1	1	14
Maria3	2;3,18	7			1		8
Maria4	2;5,4	16				3	19
Maria5	2;5,24	1	3	1			5
Maria6	2;7,1	8					8
Maria7	2;8,3	8		1	1		10
Maria8	2;8,27	4		1	2	2	9
Total		**57 (76%)**	**3 (4%)**	**3**	**5**	**7**	**75**

Table 6: Occurrences of the different types of embedded clauses in Maria's corpus.

Child/File	Age	Infinitives	That-clauses	RCs	Conditional	Purpose	Total
Sara1	2;0.20	1	1				**2**
Sara2	2;1.15	9			1		**10**
Sara3	2;2.1	14				3	**17**
Sara4	2;2.11	2		1			**3**
Sara5	2;3.2	2	2	1		3	**8**
Sara6	2;3.28	5	1			1	7
Sara7	2;4.18	3	3	6			**12**
Sara8	2;5.9	6	3	1		2	**12**
Sara9	2;6.13	10	5			2	**17**
Sara10	2;7.10	20	2	2		3	**27**
Sara11	2;8	20	4	2			**26**
Sara12	2;8.21	13	5				**18**
Total		**105 (66%)**	**29 (18%)**	**16**	**1**	**14**	**159**

Table 7: Occurrences of the different types of embedded clauses in Sara's corpus.

Child/File	Age	Infinitives	That-clauses	RCs	Conditional	Purpose	Total
Shem1	2;2.16	16					16
Shem2	2;2.23	16					16
Shem3	2;3.2	22					22
Shem4	2;3.16	16					16
Shem5	2;3.21	21		1			22
Shem6	2;3.28	7				2	9
Shem7	2;4.4	14					14
Shem8	2;4.20	9					9
Shem9	2;4.25	-					-
Shem10	2;5.2	12					12
Shem11	2;5.9	-					-
Shem12	2;5.16	35					35
Shem13	2;5.23	32					32
Shem14	2;5.30	-					-
Shem15	2;6.6	25					25
Shem16	2;6.27	31	2	3			36
Shem17	2;7.10	24		1			25
Shem18	2;7.18	28	5		1		34
Shem19	2;7.26	33	1				34
Shem20	2;8.3	-					-
Shem21	2;8.15	13		1			14
Shem22	2;8.20	23	1				24
Shem23	2;8.29	11	1				12
Total		**388 (95%)**	**10 (2%)**	**6**	**1**	**2**	407

Table 8: Occurrences of the different types of embedded clauses in Shem's corpus.

Some examples from the different types of embedded clauses produced by the three children are given below.

(4) a. thelo ego na[10] to kano, ela!
I-want, I Mood(na) Clitic/this to do, come on
I want to do this, come on!

b. i statobuta nomizi oti ine ati i statobuta
the Cinderella thinks that is this-Fem the Cinderella
Cinderella thinks that she is Cinderella

c. ... to eleje i migula xxx, pu' tane poli megali, i migula
... (it) said the fly-Dim xxx, that was very big, the fly-Dim
she said this, the little fly xxx, who was very big, the little fly

d. tha k(l)iso ego tin porta ama erthi
I-will-Fut close-Fut I the door if comes
I will close the door, if he/she comes

e. prepi na to anijis etsi ja na kani xxx
need Mood it open in this way in order to do xxx
There is need to open it in this way, in order to do xxx

[10] As mentioned previously, the verbal form "na"+verb in Modern Greek is not an infinitive, but rather a mood phrase. In fact, there is a debate open for this subject in the literature (Varlokosta, Vainnika & Rorhbacher, 1998, Hyams 2002, among others) but there is no reason to discuss that in this analysis. Consequently, I will refer to this form as the infinitive, for convenience's sake, being some kind of equivalent of the infinitive of other languages, at least for immediate translation.

Examples in (4) are reported from Maria's corpus. Each example corresponds to a different type of embedded clause as given in table 2, and hence a. is a subjunctive na-phrase (infinite), b. that-clause, c. RC, d. conditional and e. purpose sentence. It seems that in Modern Greek, as in Italian and English, the most prominent form with respect to the percentage of occurrences, is the subjunctive na-phrase, i.e. the infinitive, 76%.

(5) a. no, non ci riesco a mettere il dito dentro.
no, not (ci)Particle I-arrive to put the finger inside
I don't arrive to put the finger inside

b. sai che io mi sono b(r)ucciata il ginocchio.
you-know that I me burn-Reflexive the knee
do you know that I have burnt the knee.

c. ti ricordi quel [/] quel cane nero che mi guardava?
you-remember that (that) dog black that me looked
do you remember that black dog that was looking to me?

d. non so se viene.
no know if comes
I don't know if he/she comes

e. io so(no) così &co contenta di venire qua pel giocare
I am so happy to come here (in order) to play
I am so happy to come here, in order to play

(6) a. i wanna try something xxx the horsie #[11] kay?

b. i got it # i think that we are ready to show pancake+book to ana now # den.

c. [/] dese are the sticks (that) you drum on a drumber+thing da(t) is a drum

d. yeah # see i(f) dis broke # i wanna put some words <tuh> [/] too an(d) <no(t)> [/] an(d) not goin(g) like dis.

e. (h)e t(r)ies wi(th) dis tuh thing to take duh ball.

In (5) and (6) examples from Sara and Shem respectively represent the same pattern given in (4) previously. Infinitives, again are more prominent; 66% for Sara's speech production, and 95% reported from Shem.

3 Acquisition of Embedded Clauses –*Proposal*

The emergence and development of embedded clauses in a child's grammar is a vital milestone for its grammatical competence in that it provides evidence of the presence of the recursive rule.

[11] Pause between words

The account of early language acquisition of embedded clauses proposed in this section supports the hypothesis of continuous development, i.e., that grammatical development is constrained by principles and parameters of UG.

It also provides empirical support for the hypothesis that the child grammar of a particular language may differ from the adult grammar of the language but only minimally and in well-defined limits. Such empirical evidence, it will be given in what follows, concerning the analysis of the results from the data. More specifically, we will see soon, children's production extracted for the present study contains functional categories and embedded clauses and that such complex constructions are produced following the adult model.

The restructuring process (parameter resetting), if necessary for some of the children, will be delimited by the limits imposed by UG. As stated in Section 1, the children's grammar may be a subpart of adult's grammar, constrained by the principles and parameters of UG. In this sense, the restructuring process consists in the resetting of those parameters fixed in a first instance from children not following the adult target grammar. Thus, all the formation processes towards the final adult-like grammar and the passage from the minimally different early grammar to the Stable Stage are always subject to constraints and limits imposed by UG. Specifically in this analysis the hypothesis proposed is that the emergence of embedded clauses in children's grammar may arise early in the grammatical development of some of the target children. Moreover, such complex constructions are consistent with the adult grammar model.

The following table presents the time of emergence and first use of the different types of embedded clauses in the speech production of the three children:

First use of	Maria	Sara	Shem
Infinitive	2;0.24	2;0.20	2;2.16
That-clause	2;5,24	2;0.20	2;6.27
RCs	2;5,24	2;2.11	2;3.21
Conditional	2;2,8	2;1.15	2;7.18
Purpose	2;0.24	2;2.1	2;3.28

Table 9: Age of first use of the various types of embedded clauses in the speech production of Maria, Sara and Shem.

What is interesting to notice in this table, is that the three children in this study differ minimally in the timing of the first use of embedded clauses. In fact Maria and Sara are very close to each other with respect to the timing of emergence of most types of embedded clauses whereas Shem's average age of first use is a little bit higher.

The onset of embedded clauses is particularly interesting because Crain & Lillo-Martin (1999) claim that children begin to use multi-clause sentences, such as RCs, complement clauses, and conjoined clauses[12] at approximately in 3 ½ to 4 years of age (Stage IV), whereas children at approximately 2 ½ to 3 years of age, i.e. at Stage II, still produce sentences of just one clause.

It is very interesting that the empirical results of the present work suggest that children supposedly at Stage II (following Crain & Lillo-Martin) seem to manage and produce complex clauses. In particular, in ref. to table 9, I have shown that the emergence and consequently use of embedding is placed importantly earlier than suggested in other studies. To sum up, the main finding of the present study is that

[12] The assumption is that conjoined clauses are mastered earlier because they have flatter phrase structure than RCs and other embedded clauses.

children acquire embedded clauses at about 2½-3 years of age, depending on the experience and target circumstances.

The empirical evidence of this work supports further the claim for the early acquisition of embedded clauses, and hence that in the range of age from 2;0.20 to 2;7.18, the emergence and production of embedded clauses according to the adult model, has been completed. The result is valid for all types of embedded clauses analysed in this study.

Furthermore, it is also worth noting that the combination of results shown in ref. tables 6, 7, 8, 9 indicate that the most prominent use and production of embedded clauses concerns infinitive constructions. Interesting in all children's corpora, what emerge first and in vast proportion are infinitives: 76% out of the total constructions for Maria, 66% for Sara and 95% for Shem. Specifically, in Maria's corpus 57 instances out of 75 are infinitives, in Sara's 105 out of 159, whereas in Shem's corpus 388 out of 407. In spite of the high prominence of infinitives, the other types of embedded constructions are produced by the children as well.

Some types of clauses emerge later than others do, and interestingly, others show only a few instances in the speech production of the children. Children differ from each other due to cross-linguistic variation and factors concerning the triggering input. For example, in the corpus produced by Shem, only 1 instance of a conditional embedded clause occurs, as did Sara, whereas, in Maria's corpus we found 7 well-formed conditional constructions. Maria and Sara produce an interesting amount of purpose constructions, while in Shem's corpus only 2 instances were found, which is probably due to the fact that purpose constructions are easier in Modern Greek and Italian than in English. Compared to other types of embedded clauses, that-clauses are produced extensively in the corpora of all children. Infinitives and that-clauses, being arguments of the verbs and not adjuncts, are more prominent in the production of the three children.

Another phenomenon presented in the data, is the consistent presence and production of RCs by Sara, but not by Maria and Shem. Thus, in Sara's corpus there were 16 instances of RCs out of 159. Such variation may perhaps be due to certain structural linguistic differences in RCs in the languages studied here; as mentioned above, RCs in Italian are less complex than in the other languages studied.

It can also be noted that the ungrammatical use of embedded constructions in all children's corpora amounts to less than 1.2% of the total of embedded constructions attested. Children studied for this work not only produce embedded constructions of all the types suggested previously, the utterances produced were grammatical and followed the adult grammar model. This fact clearly shows the similarity and closeness of the children's use of embedded clauses to the adult target.

It has been suggested in various studies that children's speech is telegraphic or corresponds to small-clauses constructions of the adult target. Theories such as the Small-clause hypothesis among others suggest that children's grammar lack functional categories and consists only of lexical categories. Thus, initially, structures accommodate all the arguments, but do not include functional categories. In other words, the grammar only produces thematic structures, whereas the maturation of the functional system follows only later. Tsimpli, (1992) suggests that the total of functional categories comprise an independent functional module within the language faculty distinct from the lexical one. According to Tsimpli[13], UG principles are

[13] Quantitative data from the corpus are not presented in Timpli's study. All observations regarding the absence of the various functional categories were presented based upon anecdotal observations.

always available, but the functional module is subject to maturation, and hence initially not available. Based upon these assumptions, Tsimpli's theory arrives at the conclusion that pre-functional grammars are possible grammars in the sense defined by UG and that parameterisation is absent. The hypothesis that a process of maturation exists, having effects on language development, it's plausible. Nevertheless, the absence of functional categories in children's phrase structure, proposed by Tsimpli, has a number of consequences regarding the linguistic availability of null subjects, the absence of movement processes, the absence of case assignment , embedding etc.

However, the present study suggests that functional categories are indeed present in the children's grammatical system. Functional projections are projections of functional heads, such as CP, INFL or Det. The underlying assumption is that, like lexical heads N, A, V and P, functional heads have a syntactic projection as dictated by X-bar Theory. Children's clauses are not just lexical-thematic representations, but representations with lexical and functional projections since introductory particles for embedding are always positioned higher in the CP layer. Early emergence and use of embedded clauses studied in this analysis support the claim of the presence of a functional system in children's grammar. Additionally, studies like Borer & Wexler (1987), Radford (1990 & 1997), Varlokosta, S., Vainnika, S. & B. Rorhbacher (1998) give further evidence for the presence of the other functional projections, like IP and Det in children's grammar.

4 Conclusions

The main goal of this research was to give a proposal of acquisition of embedded clauses. Based on different theories concerning language acquisition, I have analysed data from the naturalistic speech productions of various children.

What is suggested is that children of about 2 ½ to 3 years of age, have indeed acquired the structure and distribution of embedded clauses, according to the adult target, and moreover, they produce such complex constructions.

In diverse discontinuity models, it has been suggested that children acquire complex structures later on in their grammatical development or that such constructions are subject to restructuring.

Various explanations have been proposed in order to analyse and account sufficiently for such phenomena. Bellugi (referred to in Hyams, 1986) suggests that the child's reduced responses reflect those portions of the adult sentences which the child can analyze successfully, and hence there is a filtering effect activated for the input data.

It is reasonable to expect then, that those elements, which are not analyzed by a particular grammar, will be neither systematically interpreted nor produced by the speaker/hearer. With regard to this last property it is proposed that the early grammar exerts a filtering effect on the data. The child systematically ignores a class of elements for which his grammar provides no structural description.

The question that arises from such a claim is how do children begin to take into consideration items that they previously ignored? Children, as noted previously, may learn the lexicon on a word-to-word basis, hence the properties and subcategorization frames of each verb as each verb is encountered.

In the same spirit, Hyams (1985) has suggested that children do not acquire sentential complementation until they learn that verbs can take propositions as arguments. On this analysis the development of a particular syntactic ability is

triggered by the development in the semantic/conceptual domain and markedness is not a relevant factor. For example, it may be the case that embedding is introduced into the grammar by the acquisition of verbs which subcategorize a sentential complement (Hyams, 1985). In fact, complements subcategorized by a given verb will appear within a month after the first use of that verb in any construction. The first complex sentences to emerge contain verbs such as "want" and "watch" which children previously use in conjoined and "flatter" clauses. Thus, the acquisition of particular verbs induces a restructuring in the grammar to include sentential embedding under VP.

More generally, it has been suggested that grammatical development is a "continuous" process in the sense that each of the intermediate grammars falls within the limits imposed by UG. On a continuous model, however, the restructuring which one expects is within narrowly defined limits.

What I found from this research is that children in between 2 ½ and 3 years of age make extensive use of complex sentences and moreover, such structures correspond to the adult target model. Verbal arguments emerge earlier and in higher proportions than adjunct embedded material, and functional categories (in particular CP layer) are present within the syntactic structure of such constructions in children's grammar.

The acquisition of the syntactic system does not involve any complex learning mechanisms. Rather, it is biologically determined to mature at a particular point in development. Thus, the child begins with a grammatical system perhaps different from the adult one and undergoes a kind of metamorphosis, the output of which is an adult-like syntactic system.

From these findings we have inferred that children's representation of embedded clauses does not encode only lexical information, but also functional information, syntactically expressed through IP and CP layers of a clause.

The data at our disposal is enough to reject the claim of the various hypotheses against maturation and continuity, according to which functional categories and children's grammar in general are subject to a restructuring process. The cross-linguistic investigation of the present study makes plausible and gives evidence that there is early availability of functional categories.

References

Agouraki, Y. (1991). "A Modern Greek complementiser and its significance for UG". UCLWPL, Volume 3

Bauer, B. (1995). "The emergence and development of SVO patterning in Latin and French". New York: Oxford University Press.

Bernstein, J. (1993). "Topics in the Syntax of Nominal Structure Across Romance." PhD Dissertation, The City University of New York.

Borer, H. & Wexler, K. (1987). "The Maturation of syntax". In Roeper & Williams (eds.) *Parameter Setting.* Dordrecht: Reidel, 1987.

Brain, D. J. (2000). "On Some Control Structures in Hellenistic Greek: A Comparison with Classical and Modern Greek" The Ohio State University

Cardinaletti, A. & M.T. Guasti. (1995). "Small Clauses. Syntax and Semantic." U.K.: Academic Press.

Chomsky, N. (1980). "On Binding". Linguistic Inquiry, 11

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

Chomsky, N. (1986). "Barriers." Cambridge, Mass.: MIT Press.

Chomsky, N. (1989). "Some notes on Economy of Derivation and Representation." In Laka and Mahajan (1989)
Chomsky, N. (1995). "The minimalist program." MIT Press.
Cook, V. J. (1990). "La grammatica universale. Introduzione a Chomsky. (Universal Grammar; Introduction to Chomsky)" Bologna: Il Mulino
Cinque, G. (1999). "Adverbs and Functional Heads: A cross-linguistic perspective." Oxford University Press.
Crain, S and Pietroski, (2001). "Nature, Nurture and Universal Grammar" Linguistics and Philosophy 24: 139–186. Kluwer, Academic Publishers. NL.
Crain, S. and Lillo-Martin, D. (1999). "An Introduction to Linguistic Theory and Language Acquisition". Blackwell Publishers Inc.
Crain, S. and R. Thornton, (1998). "Investigations in Universal Grammar. A Guide to Experiments on the Acquisition of Syntax and Semantics." MIT Press
Doukas, T. (1998). "The null subject parameter and non-finite forms on native Greek language acquisition. Study of a case". BA Dissertation, University of Sienna.
Guasti, M. T. (2002). "Language Acquisition. The Growth of Grammar." MIT Press
Guasti, M. T. (1994). "Verb syntax in Italian child grammar: Finite and non-finite verbs". Language Acquisition 3(1): 1-40.
Guasti, M. T. (1993). "Causative and Perception Verbs. A Comparative Study." Torino: Rosenberg & Sellier
Haegeman, L. (1994). "Government and Binding Theory". Blackwell Publishers Ltd.
Haegeman, L. (1996). "Manuale di grammatica generativa (Handbook of Generative Grammar)". Milano: Ulrico Hoepli Editore S.p.A.
Haegeman, L & J. Gueron, (1999). "English Grammar; A Generative Perspective". Blackwell Publishers Ltd.
Haegeman, L. (2002). "Speculations on adjunct and the left periphery." Seminar notes at SOAS, March 2002.
Hyams, N. (2002). "Clausal Structure in Child Greek: A reply to Varlokosta, Vainikka and Rohrbacher and a reanalysis".
Hyams, N. (1986). "Language acquisition and the theory of parameters." Reidel Publishing Company, Kluwer Academic Publishers Group.
Hyams, N. (1985). "The Acquisition of Clausal Complementation". Paper appeared in the Proceedings of WCCFL IV, Stanford Linguistics Association, 1985
Ingram, D. (1989). "First language acquisition. Method, description and explanation". Cambridge University Press, 476-483.
Katis, D. (1981). "I kataktisi tu neo-elliniku rimatos apo to pedhi (The Acquisition of the Modern Greek Verb)" Studies in Greek Linguistics 2, Thessaloniki: 269-289
Katis, D. (1985). "I morphologiki grammatiki sta proima stadia tis pedhikis glosas (Morphological grammar in the early stages of child language)." Studies in Greek Linguistics 6, Thessaloniki: 115-129
Lightfoot, D. (1991). "How to set Parameters: Arguments from Language Change." MIT Press.
Marinis, T. (2000). "Acquiring the Left Periphery of the Modern Greek DP" Paper presented in the Peripheries Conference, at York, in September 2000.
O'Grady, W., Y. Yamashita, M. Lee, M. Choo & S. Cho. (2000). "Computational factors in the acquisition of relative clauses". Talk at International Conference on the Development of the Mind, Keio University, Tokyo, August 2000.
Philippaki Warburton, I. (1998). "Functional categories and modern Greek syntax." In Drachman, G. & Theophanopoulou-Kontou, D. (eds.). *Linguistic Review* 15: 159-186.

Poeppel, D. & K. Wexler. (1993). "The full competence hypothesis of clause structure in early German." Language, 69:1, 1-33.
Poletto, C. (2002). "The Higher Functional Field. Evidence from Northern Italian Dialects." Oxford University Press.
Pollock, J-Y. (1989). "Verb movement, Universal Grammar and the Structure of IP." Linguistic Inquiry 20: 365-424.
Puskas, G. (1997). "Focus and the CP domain." in L. Haegeman (ed.) "The new Comparative Syntax." (1997).
Radford, A. (1990). "Syntactic Theory and The Acquisition Of English Syntax". Basil Blackwell, Ltd.
Radford, A. (1997). "Syntactic theory and the structure of English; A minimalist approach". Cambridge University Press.
Rizzi, L. (1982). "Issues in Italian Syntax." Dordrecht: Foris
Rizzi, L. (1990). "Relativized minimality." Cambridge, Mass.: MIT Press.
Rizzi, L. (1992). "Early null subject and root null subject." Geneva Generative Papers, vol. 0, no. 1-2.
Rizzi, L. (1993/1994). "Some notes on linguistic theory and Language Development: The Case of Root infinitives." Language Acquisition, 3(4): 371-393.
Rizzi, L. (1997). "The fine structure of the Left Periphery". In: L. Haegeman (ed.) Elements of Grammar. Handbook in Generative Syntax. Dordrecht:Kluwer, 281-337.
Rizzi, L. (2001). "Locality and Left Periphery." Universita Degli Studi di Siena
Roussou, A. (2000). "On the left periphery: modal particles and complementisers" Journal of Greek Linguistics, Volume 1 (2000), pp. 65-94
Schutze, C. T. e K. Wexler. (1996). "Subject Case Licensing and English Root Infinitives." Cambridge, Mass.: MIT Press.
Tsimpli, I. M. (1992). "Functional Categories and Maturation: The Prefunctional Stage of Language Acquisition." Ph.D. Dissertation UCL, London.
Valian, V. (1991). "Syntactic subjects in the early speech of American and Italian children." Cognition, 40: 21 - 81.
Varlokosta, S. (1999). "The structure of prepositional relatives in Modern Greek: wh-movement or pronominal linkage?". In A. Mozer (ed.), Proceedings of the 3rd International Conference on Greek Linguistics, University of Athens.
Varlokosta, S., Vainnika, S. & B. Rorhbacher. (1998). "Functional projections, Markedness, and Root Infinitives in early child Greek". The Linguistic Review 15: 187-207
Varlokosta, S. & S. Armon-Lotem (1998). "Resumptive and Wh-movement in the Acquisition of relative clauses in Modern Greek and Hebrew". Proceedings 22nd Boston University Conference on Language Development (Cascadilla Press).
Wexler, K & M. R. Manzini (1985). "Parameters and Learnability in Binding Theory". In Roeper, T & E. Williams (eds.) *Parameters and Acquisition Theory*
Wexler, K (1992). "Some Issues in the Growth of Control". In R. K.Larson, S. Iatridou, U. Lahiri, & J. Higginbotham, (eds.) *Control and Grammar*. Dordrecht: Kluwer.

Other sources

http://users.info.unicaen.fr/~tlebarbe/Linguistics_Lexicon/ll_a.html
http://www.google.com
http://linguistlist.org/
http://www.linguistlistplus.com/