THE UNIVERSITY *of* EDINBURGH

The Emergence of Functional Categories in Bilingual First Language Acquisition

Ludovica Serratrice

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy to the

University of Edinburgh

December 1999

Abstract

This thesis is a case study on the emergence of functional categories in bilingual first language acquisition. The investigation focuses on the transition from one-word to multiword utterances and the shaping of functional projections of Determiner, Agreement and Tense and their associated formal features.

The empirical basis of this work is a corpus of thirty-nine videorecorded observations of Carlo, an English-Italian bilingual child, during free-play sessions with an adult. Data was collected separately for English and Italian for a period of fifteen months from when the child was 1;10 until he was 3;1, and was then transcribed in CHAT format.

Four interrelated lines of enquiry inform the analysis presented here. The principal research question concerns the acquisitional strategies adopted by C. in these early stages of development in the two languages. A bilingual child is the closest one can get to a perfect matched pair where a number of variables such as socio-cognitive development, socio-economic status, parents' education, etc. are eliminated, and the two main variables to be investigated are the child's two input languages. This is an ideal situation in which the respective roles of general acquisitional strategies and language-particular ones can be teased apart. An analysis of the emergence of the morphosyntactic correlates of Determiner, Agreement and Tense categories in English and Italian reveals a discrepancy between the two languages in the age of acquisition, rate of acquisition and in the language-specific strategies the child adopts.

The observation of a significant difference in C.'s acquisitional strategies in English and Italian leads us to the second and third research questions: the way in which the emergence of functional categories differs between the two languages, and the reasons why this should be the case. The most obvious difference is the extent to which morphological correlates of functional categories emerge in the child's speech. In Italian, verbal and nominal morphology emerges earlier than in English and, at least in the nominal system, there is evidence that an Agreement category is part of the child's grammar. In English, verbal morphology is virtually non-existent by the end of

the period of observation, and there is no substantial evidence that either Agreement or Tense are realised. Lexically-specific, item-based learning plays a substantial role in both languages, but in Italian there is some evidence that a number of grammatical contrasts are becoming productive by age 3;0, albeit some of them are still limited to a small number of lexical items.

Two reasons were identified for the observed differences in the emergence of Determiner, Agreement and Tense in English and Italian: a typological reason, and an environmental reason. The former concerns the richness of Italian morphology, where grammatical contrasts are transparently marked both on nominal and verbal paradigms, as opposed to the relative poverty of English morphology where such contrasts correlate less obviously ans systematically with morphophonological markers. The latter reason concerns the very different input conditions in which C. is exposed to Italian and English: Italian is the home language spoken to him by his family and his babsysitters, while he is addressed in English by the staff at the nursery where one adult is in charge of several children and cannot engage in the one-to-one interaction which is typical of the dyadic situation in which C. finds himself at home.

The differences observed in the lead-lag pattern between C.'s Italian and his English also provide sufficient evidence to address the fourth research question concerning the separate developement of the two languages. The analysis of the data did not reveal any systematic interferences from one language to the other. On the contrary there is evidence that C. is sensitive to the different morphosyntactic cues of his two input languages, and that he can treat the two as independent, self-contained problem spaces.

Acknowledgements

My first debt of gratitude is for my two supervisors: Ronnie Cann and Antonella Sorace. Thank you Ronnie for your patience, for your advice and support over the last four years, and for helping me develop my ideas without ever imposing yours. Antonella, thank you for your interest in my work, for your precious criticism, and for reminding me what a PhD is all about (including weekly tennis practice!).

A number of people have helped me to collect the data for this thesis, and I am extremely grateful to all of them. First and foremost Carlo and his family. Carlo, you have been a constant source of amazement over the last four years, and it goes without saying that there is a lot more to your language acquisition than I could possibly do justice to in this thesis. I hope you will enjoy reading it when you are a little older. Thank you to Carlo's parents, Antonella Sorace and Bob Ladd, for always supporting this project so enthusiastically; to Marco Ladd for his cooperation in some of the data collection and for being an amazing child; to Andrew Ladd for putting up with occasional disruptions to his TV viewing schedule. Two people have been especially generous with their time and have made the collection of English data possible: Karen Kay and Eric Laurier. Thank you Karen for spending many hours playing with coloured wooden blocks, and thank you Eric for your expertise with Lego and picture domino. Thanks also to Elena, Alessandra, Claudia, Bob, Jim, Struan, Sarah, Amy, Rachel, Claudia and Josh for participating to some of the recording sessions. I am also grateful to the children and staff of Forbes Nursery and Edinburgh Kindergarten for allowing me to videotape them on a number of occasions. John Glendinning and Barbara Brown at the Language and Humanities Centre have also been instrumental in the data collection process by making available various bits of equipment, often at the last minute.

Completing this thesis would not have been possible without the financial support of a University of Edinburgh Postgraduate Studentship, for which I am very grateful. At the University of Edinburgh, a big thank you to Jim Miller for having been a very understanding head of department during my years in the Department of Linguistics, and for exchanging papers and ideas on language acquisition. Thanks also to John Joseph for allowing me access to the facilities in the Department of Applied Linguistics at a time when I really needed it. Technical staff have also been

very generous, thanks to Irene McLeod for unfailing kindness, and to Eddie Dubourg for always knowing the right answer to everything. I would also like to thank Ethel Jack for being there in the beginning, and for saving me from the photocopying machine more times than I care to remember, and Fiona Paterson for being there in my last year and for always doing everything so competently and cheerfully.

My colleagues and friends in the Department of Linguistics and in the Department of Applied Linguistics deserve a special thank you: Iraide Ibarretxe for her spider plants, and for teaching me something on perception verbs and Basque (Irune etorri da); Aaron Drews for making me aware of intrusive r, and for being a Mac fan; Anna Babarczy for questioning some of my assumptions on language acquisition, and for explaining chi square; Matthew Bull for introducing me to film-going in Edinburgh; Max Louwerse for still being able to joke at 8.30 p.m. on Fridays in the computer lab; Satu Manninen and Kook-Hee Gill for putting up with my boxes in their office; Szilvia (Sz) Papp for setting up the language acquisition reading group together, and for friendly advice; Charlie Kemp for always being so calm and reassuring; Sonia Rocca for sharing stories about working with children; and Patricia Soley and Bella Vivat for late-night dinners and great fun.

A very special acknowledgement goes to Eric Laurier for everything in the last seven years: for discussion about linguistics and the merits of ethnomethodology (and many other things), for letting me take care of Spunky and Lola, and for being my friend.

Finally my deepest thanks to my parents, Jose and Piermario Serratrice, for trusting me and supporting my choices in every possible way with extreme generosity and altruism; to my sister Laura for being the best sister ever; and to Massimo Mazzotti for doing more than I can thank him for. Grazie!

Table of contents


CONIGLIETTO
ROSA
BUNNY
ROSE
MARGHERITA
GATTO
CAT
CARLO
GIRASOLE
SUNFLOWER

Chapter 1

Introduction

1.1. The ontology of Functional Categories in child language

This thesis is a study of the early stages of linguistic development of Carlo (C.), an English-Italian bilingual child, between the ages of 1;10 to 3;1. The focus is on the transition from one-word to multiword utterances and on the emergence of the morphosyntactic correlates of Functional Categories (FCs). FCs are intended here as the abstract representation of formal grammatical features encoded by lexical items in a speaker's mental lexicon (Chomsky, 1995).

The ontological status of FCs in child language has been at the top of the research agenda in generative approaches to language acquisition for a number of years now, and various proposals have emerged in the literature. Some scholars have claimed that children's initial grammatical representations lack FCs altogether. The initial state of first language acquisition is thus a prefunctional stage where FCs are not accessible to the child; only subsequently do they become available as the result of biological maturation (Lebeaux, 1987; Guilfoyle & Noonan, 1988; Radford, 1990, 1992, 1994, 1996; Platzack, 1990, 1992; Tsimpli, 1996).

This notion of a two-tier system where lexical categories such as Noun, Verb, Adjective, and Preposition are part of the child's grammar, while FCs such as Agreement, Tense or Determiner are not, has been attacked both on theoretical and empirical grounds by a number of researchers who argue for the *a priori* availability of FCs (Pinker, 1984; Hyams, 1986, 1992a, 1992b, 1996; Verrips & Weissenborn, 1992; Poeppel & Wexler, 1993; Deprez & Pierce, 1993; Hoekstra & Hyams, 1995, 1998). Their principal theoretical objection is that a lexical-only child grammar creates a serious problem of discontinuity with the adult grammar. In their view the maturation argument proposed by prefunctionalist accounts does not offer an explanatorily adequate solution to the problem. Moreover, a growing number of crosslinguistic studies in languages other than English have proved that the prefunctional stage is not

universally observed in the earliest stages of language production. In fact, there is evidence that children learning languages with richer morphological systems than English use a number of functional morphosyntactic markers from very early on.[1]

However, although the discontinuity problem is eliminated if one assumes that FCs are an integral part of children's early grammar, there remains the non-trivial problem of how to account for the fact that children's linguistic production is certainly not like adults'. One of the solutions that have been proposed is known as the Underspecification Hypothesis (Hyams, 1992a) and states that, although FCs are indeed present in early child grammar, nevertheless they may be underspecified with respect to the formal grammatical features of the adult target. In other words, even if one assumes that, for example, an AGR category is present in the verbal domain of a child's grammar, at the same time it is reasonable to assume that one of the features associated with this node, namely Number, may be initially underspecified. Children for whom the feature Number is underspecified will omit inflectional Number markers and produce non-finite forms instead of finite forms as required by the target grammar where AGR must be specified for both Person and Number features (Hoekstra & Hyams, 1995; Hoekstra, Hyams & Becker, 1996).

An alternative hypothesis on the availability of FCs in early child grammar is proposed by Clahsen and colleagues (Clahsen & Penke, 1992; Clahsen, Penke & Parodi, 1994; Clahsen, 1996). In what they call a Weak Continuity approach, they concede that FCs are not subject to maturational constraints as proposed by Radford (1990) and others, but are *in principle* always available to the first language learner. Differently from advocates of a stronger continuity approach however, they argue for the lexically-driven instantiation of FCs. A FC can be said to be part of a child's grammar only when there is morphosyntactic evidence for its existence.

The main difference between a Weak Continuity approach and a Strong Continuity approach lies in the way in which the two hypotheses view the relationship between FCs and their overt realisation through morphophonological markers. For Weak Continuity approaches the productive use of a given set of morphosyntactic markers implies the existence of the relevant FC; in the absence of such overt lexical evidence one is to conclude that the corresponding FC is not active in the child grammar yet. For a Strong Continuity approach the relationship between FCs and their

[1] See for example Hyams (1986) and Guasti (1993/94) for the acquisition of the Agreement system in Italian, Levy (1983) for the acquisition of gender in German, Russian, Polish, French and Hebrew, Demuth (1992) for the acquisition of passives in Sesotho.

overt morphosyntactic correlates is not as straightforward. On the one hand, the early production of grammatical markers in inflectionally rich languages has been used to argue against a universal prefunctional stage in early child language. On the other hand, from a Strong Continuity point of view the absence of such markers is not to be taken as conclusive evidence that the FC in question is not realised in the child's grammar, it simply means that the relevant morphological paradigm has not been learnt yet. While in a Weak Continuity approach the stipulation of a functional node in the child's syntactic representation is justified by empirical evidence that the corresponding formal grammatical features have been realised by the child, in a Strong Continuity approach the existence of a FC is an *a priori* theoretical requirement.

The sharp dissociation between lexical and grammatical learning that is at the core of some of the strongest continuity arguments may have to undergo some serious revision in the light both of recent developments in syntactic theory where the lexicon occupies an ever more central role (Bresnan, 1982, in press; Pollard & Sag, 1994; Goldberg, 1995; Chomsky, 1995), and on account of an expanding body of research uncovering the powerful relationship between lexical and grammatical growth (Bates, Bretherton & Snyder, 1988; Marchman & Bates, 1994; Bates & Goodman, 1997, 1999; Goodman, 1995; Jahn-Samilo, 1995; Caselli, Casadio & Bates, 1999). One of the findings from the Bates, Bretherton & Snyder (1988) study, for example, is that the best predictor of grammatical development at 28 months (in the heart of the "grammar burst") as measured by MLU, is vocabulary size at 20 months (in the heart of the "vocabulary spurt"). In another large-scale study of the lexical and grammatical abilities of English-speaking children between 16 and 30 months, Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly (1994) found a strong correlation between grammatical complexity, as measured by the 37-item sentence complexity scale on the Toddler version of the MacArthur Communicative Development Inventory, and productive vocabulary size. A similar relationship between vocabulary size and grammatical development has also been reported for two large groups of Italian- and English-speaking children between 1;6 and 2;6 by Caselli, Casadio & Bates (1999), despite striking differences between the two languages in the content of vocabulary and grammar.

The central theoretical assumption subscribed to in this work is that syntactic development is driven by lexical learning. It is justifiable to speak of the realisation of a FC only if and when there is convincing empirical evidence that the corresponding morphosyntactic markers are being used by the child in a productive and contrastive

way in at least 90% of obligatory contexts. This theoretical position is motivated by a principle of structural economy whereby it is unnecessary, and hence not desirable, to postulate more structure than is needed to account for the data (Grimshaw, 1994; Clahsen, Penke & Parodi, 1993/94). If there is no overt morphological and/or syntactic evidence that a child is sensitive to a Tense or an Agreement contrast, for example, there is no need to assume a vacuous T or AGR projection that is neither associated with morphological material, nor is it the landing site of a syntactic movement operation.

1.2. Abstract categories and lexically-driven positional patterns

From a more specific methodological point of view we also believe that it is necessary to further qualify what kind of empirical evidence justifies the postulation of a FC. It is often standardly assumed that a given morpheme is acquired when it is correctly produced in 90% or more of contexts in which it is obligatorily required by the target grammar (Brown, 1973). Frozen phrases or stereotyped formulas where it is clear that the child is simply using an unanalysed string are normally eliminated from the count of productivity. Deciding what counts as a formula in child language is however not always an easy task. Several definitions of formulas are available in the literature (Peters, 1983; Hickey, 1993; Plunkett, 1993; Pine & Lieven, 1993), and the criteria for identifying them include prosody, length of unit, frequency of occurrence and appropriateness of use. Formulas are normally associated with the idea of prefabricated invariable phrases such as *bye bye, oh dear, that's mine, what's that?, I don't know*, etc. of which one typically finds frequent examples in early child language. In addition to this type of more easily identifiable formulaic utterances in which lexical choices are extremely stereotyped, some researchers have argued for the existence of other phenomena of lexical specificity in early child language which may be less salient, but offer a wealth of information on the child's strategy in the acquisition of the building blocks of her language.

Capitalising on work by Braine (1976), Lieven, Pine and colleagues (Lieven, Pine & Dresner-Barnes 1992; Pine & Lieven, 1993; Lieven, Pine & Baldwin, 1997) have rediscovered the notion of positionally productive patterns where multiword utterances are unalysed in terms of frames with slots filled by constant and variable lexical items. Their working assumption is that a careful analysis of patterns in early child language reveals a high degree of lexical specificity whereby children's knowledge is best characterised in terms of limited scope formulas rather than abstract categories. Their coding scheme for child language analysis includes: *frozen phrases*

containing two or more words that have not previously occurred on their own in the child's production, or if they have, not in the same position inside the multiword utterance. *Intermediate utterances*, multiword utterances where each of the single words has occurred previously in the child vocabulary but not in the same position in two previous multiword utterances. And finally *constructed utterances,* multiword utterances where one or more words have occurred independently in the child's vocabulary before, and one word has occurred in the same position in at least two previous multiword utterances.

Using this type of coding scheme in a in-depth study of the syntactic development of 11 English-speaking children between 1;0 and 3;0, Lieven, Pine & Baldwin (1997) report that a mean of 60% of all the children's multiword utterances can be accounted for by a lexically-based positional analysis, and that the majority of all other utterances can be defined as frozen phrases. If these results are anything to go by, there is reason to believe that lexical specificity is an important phenomenon in early child language which researchers must take into careful consideration before making far-reaching claims about the acquisition of grammatical markers and the instantiation of the corresponding FCs. More stringent measures of productivity than the 90% correct tokens in obligatory contexts are needed if one is to do justice to child language data.

In this work we have used methodological tools previously adopted by Pizzuto & Caselli (1992, 1994) and Gathercole, Sebastián & Soto (1999, in press). In these studies the authors go back to some productivity criteria originally proposed for English by Cazden (1968), and apply them to the analysis of Italian and Spanish child language respectively. Pizzuto & Caselli (1992) distinguish between "first appearance" and "point of acquisition" for a given morpheme, which is defined as the first of three consecutive samples where the morpheme is produced correctly in at least 90% of obligatory contexts. Two additional criteria are also used in determining point of acquisition: firstly, each of the samples that are crucial for scoring the acquisition point must contain at least five obligatory contexts of use. Secondly, with specific reference to verbal inflections, an inflection is considered to be used productively if it appears with at least two different verbs in each of the samples, and if each of these two verbs also appears with another inflection in each of the samples. The fact that a given inflection must be used with at least two different verb types ensures some degree of productivity. At the same time, the requirement that the verb that appears with the inflection in question also appears with another inflection gives credit to the contrastive

use of the morpheme. Contrast is essential to determine to what extent the child is actually relating different forms to a paradigm, as opposed to learning them as isolated lexical items.

In the analysis of Carlo's data we have adopted these stringent measures of productive and contrastive use in the belief that this would provide a more fine-tuned evaluation of the child's acquisition path, and that it would also allow us to uncover patterns that the criterion of 90% correct production in obligatory contexts could not detect.

In the next section we briefly turn to the significance of bilingualism studies in first language acquisition research and to the research questions addressed in the bilingual case study in the present work.

1.3. Bilingualism and first language acquisition

In recent years there has been a spate of studies on childhood bilingualism and a growing interest in the empirical and theoretical contribution that research in this field can make to the study of first language acquisition in general. Bilingual children make ideal informants for crosslinguistic research in that a number of factors that may affect the process of acquisition, such as cognitive development and personality, are controlled in a bilingual situation (Meisel, 1990; De Houwer, 1990). In the absence of extra confounding variables, researchers studying bilingual children can address the central issue of the relative weight of language-specific vs. universal factors in acquisition. Indeed, a number of studies of bilingual first language acquisition have reported significant language-specific differences in the acquisition process that are clearly shaped by the input the child is exposed to rather than by universal mechanisms (Ingram, 1981; De Houwer, 1990; Meisel, 1990, 1994).

The typological nature of the input has been shown to have significant implications for the acquisition process. For example, in a study of a German-English child and a Latvian-English child, Sinka & Schelletter (1998) report that the acquisition of verbal inflections takes place earlier and faster in morphologically rich languages like German and Latvian when compared to a morphologically poorer language like English. Similar findings are reported by Paradis & Genesee (1996, 1997) for a number of English-French bilinguals, and by Gawlitzek-Maiwald & Tracy

(1996) for a bilingual English-German girl. A lead-lag pattern in the acquisition of morphosyntax is observed between languages where cues to morphological paradigms are highly transparent and reliable, and languages where morphological markers are less salient and less easily incorporated in a regular paradigm.[2]

In addition to typological differences in the nature of the input the bilingual child is exposed to, it is equally as important to identify the various social contexts in which the child hears her two languages. There is obviously a large number of possible input conditions where the child may hear her two languages in a variety of different situations from a number of people including parents, siblings, grandparents, family friends, childminders, nursery staff, etc.. Because the possible bilingual environments may vary substantially from one situation to another, it is of paramount importance that researchers clearly identify what the exposure patterns are for the children studied. To date not many studies of early childhood bilingualism have provided detailed descriptions of children's linguistic environments, but this kind of information is crucial to understand how and why a bilingual child acquires his two languages the way he does. Core issues in bilingual acquisition research such as language mixing and language separation, language choice, and comparison with monolingual children need to be studied in context and require that input conditions be analysed very carefully (see De Houwer, 1990; Lanza, 1992, 1997; Quay, 1995).

The investigation of input-driven language-specificity in the acquisition process is central in this work. The next section illustrates more in detail the research questions that have motivated this study.

1.4. Four research questions

In the present study the focus is on morphosyntactic differentiation in the early stages of language production of an English-Italian bilingual child (1;10-3;1). The emergence of morphosyntactic correlates of FCs in the nominal and in the verbal domain is analysed with specific reference to the acquisition of definite and indefinite articles, and the emergence of Agreement and Tense contrasts on verbs. Four main research questions are at the basis of this investigation:

---

[2] See for example the obvious difference between the Italian present tense paradigm where each person/number combination is identified by a unique inflectional ending, and the English present tense paradigm where only singular number is marked by *-s* on 3 p.s. forms.

(1) What is the child's path to the acquisition of formal grammatical features in English and in Italian?

(2) To what extent does the emergence of formal grammatical features differ in the two languages?

(3) If there is an observed difference between English and Italian, what factors can be identified to account for it?

(4) Is there evidence for the separate and independent development of the two languages?

The first task is to describe the data to evaluate the profile of the child's grammatical development. Both quantitative and qualitative criteria are used to assess the child's progress at different chronological and MLUw stages. The second and the fourth questions address the central issue of language-specific differences. With respect to the acquisition of grammatical contrasts of Tense and Agreement, and the emergence of formal Gender and Number features in the nominal systems, together with the notion of referentiality, we want to find out whether there is any evidence for a developmental asynchrony between the two languages. Moreover we are interested in the shaping of language-specific acquisitional strategies, and in the extent to which C. proves to be sensitive to language-particular properties of the input. Proof that the child is consistent in the use of English morphosyntactic markers in his English-only utterances, and in the use of Italian morphosyntactic markers in his Italian-only utterances would support the hypothesis that the two languages are developing as two separate and independent systems.

Ultimately, the goal is to find explanations for observed behaviours and our task here is to account for C.'s acquisition of English and Italian, and any similarities and differences that we may observe. We believe that attention needs to be paid both to the typological nature of the input the child is exposed to, and to the social contexts in which he hears his two languages. A careful analysis of the child's linguistic environment is in order if one wants to disentangle the various factors that may affect the course of acquisition.

The following section provides an outline of this thesis with a brief description of the contents of each chapter.

1.5. Outline of thesis

The nine chapters that follow this introduction are divided into two main sections. The first half including chapters 2, 3, 4, and 5 reviews theoretical issues in language acquisition and syntactic theory, while chapters 6, 7, 8, and 9 cover the analysis and the discussion of the empirical findings. Chapter 10 presents concluding remarks and directions for further research.

Chapter 2 reviews the state of the art in childhood bilingualism studies with particular reference to the debate on language differentiation. Issues of language choice, parental input, discourse strategies, and pragmatic and morphosyntactic differentiation are discussed in relation to the central issue of the separate and independent development of two languages in cases of bilingual first language acquisition.

Chapter 3 explores definitions of bilingual environments and introduces Carlo, the informant of this case study. Detailed information is provided on patterns of language exposure, and features of the Italian and of the English input the child recives on a daily basis. A description of the methodology of data collection and transcription is also presented in Chapter 3 and supplemented by Appendix 1 and Appendix 2.

Chapter 4 introduces elements of the Minimalist Program (Chomsky, 1995), the theoretical framework adopted in this work, and reviews a number of approaches to language acquisition research.

Chapter 5 is a theoretical introductory chapter to the syntax, semantics and pragmatics of Determiner Phrases (DPs) with specific reference to definite and indefinite articles in English and Italian. A selected review of previous studies of the acquisition of articles is also provided, together with references to the acquisition of Number and Gender marking in the nominal system.

Chapter 6 is devoted to the analysis and discussion of C.'s acquisition of definite and indefinite articles in English and in Italian, and the realisation of higher functional projections inside the DP.

Chapter 7 reports findings on C.'s use of verb forms in English. A detailed discussion is provided on the emergence of the copula, aspectual auxiliary *be,* modal verbs, and auxiliary *do*. The use of overt and null subjects is also discussed at some length.

Chapter 8 presents the results of C.'s acquisition of Italian verbal morphology. The application of stringent criteria of productive and contrastive use to the child's production reveals a verb-based acquisitional strategy where new inflections are learnt and used only with a restricted number of verb types.

Chapter 9 is an in-depth comparison of previous findings on the development of C.'s morphosyntax in English and Italian in the nominal and verbal domain. The final section is devoted to the marking of Person deixis in the two languages, and to the significance of overt subjects in English in the absence of morphosyntactic correlates of finiteness.

Chapter 10 summarises the main findings of the case study and and provides some suggestions for further research.

Chapter 2

Methodological and theoretical issues in bilingual first language acquisition

2.1. Terminology

From a developmental perspective, the simultaneous acquisition of two (or more) languages is of particular interest in determining to what extent acquisition itself is a function of the grammatical system being acquired, as opposed to the specific characteristics of the individual or of the environment. The conflation of two different linguistic inputs in the same individual provides a privileged standpoint for the researcher: the acquisition process can be investigated crosslinguistically in one single individual with the obvious advantage of eliminating variables such as cognitive and social development, socio-economic status, etc. The bilingual child is his or her own "perfect matched pair" (De Houwer, 1990: 1)

In consideration of the twofold task which confronts the bilingual child in her linguistic development, studies of bilingual acquisition have largely focused on the controversial nature of the child's language system(s). The majority of empirical investigations of bilingual development have found varying degrees of language contact between the children's two languages. Defining exactly what form language contact takes at different stages of development and in different children is not an easy task, and different researchers have employed different terms for similar phenomena thus leading to a rather confusing use of terminology in the literature. Before moving on to consider the issue of whether the children's two languages are initially conflated into one system, or whether they are differentiated from the start, it will therefore be necessary to clarify the terminology which will be used in this work.

Volterra & Taeschner (1978) identify an initial period characterised by *lexical mixing* when words from both languages are part of one single lexicon (stage I), and they define *syntactic mixing* as a situation in which the same set of syntactic rules is

applied to two separate lexicons (stage II). Redlinger & Park (1980:337) report "an initial mixed stage in language production consisting of indiscriminate combinations of elements from each language". In a study on the bilingual acquisition of English and Estonian by her son Raivo, Vihman (1985) points out a qualitative difference between infant language mixing before the dawning of metalinguistic awareness and consequent language differentiation, and a *code-switching* strategy emerging later on (around the age of 5) which is syntactically and pragmatically constrained. Genesee (1989: 162) uses mixing to define "interactions between the bilingual child's developing language systems" and code switching to refer to a "sophisticated, rule-governed communicative device used by linguistically competent bilinguals" (1989:164).

It is generally claimed that an essential pre-requisite for code-switching is the speaker's awareness that s/he is dealing with two separate linguistic systems.[1] The implication for the issue of whether language mixing in young bilingual children can be considered as an instance of code-switching or not revolves around the other crucial issue of language separation. Meisel (1989, 1994) claims that mixing might result from a pragmatic and/or grammatical deficit, thus the fact that a child mixes her languages is not necessarily a reflection of an indifferentaited grammatical system. In actual fact several researchers (Genesee, Nicoladis & Paradis, 1995; Genesee, Boivin & Nicoladis, 1996; Nicoladis, 1994; Goodz, 1994; De Houwer, 1990) have provided convincing evidence that very young bilingual children (some as young as 2) have already developed some metalinguistic and sociolinguistic awareness that they are dealing with more that one language.

2.2. One system or two?

2.2.1. The one-system hypothesis

The fact that the majority of children acquiring two languages simultaneously show an initial stage in their development when they seem to mix elements from the two languages indiscriminately, has led many researchers to conclude that they must go through a phase in which the boundaries between the two languages are blurred, perhaps non-existent. According to this approach children are thought to acquire two

[1] This thesis does not explicitly analyse mixed utterances as such. The focus here is on the emergence of Functional Categories and on the issue of language separation in the syntactic domain. Following Meisel (1989) and De Houwer (1990, 1994) only utterances that contain lexical items from only one language are considered for the purpose of the present study. The terms lexical mixing and code switching, whenever used in this work, broadly refer to phenomena of language contact.

languages as a single system, without realising that they are in fact dealing with two distinct linguistic entities which have separate phonological, morphological, syntactic and pragmatic subsystems. Proponents of the unitary language system argue that the transition to a gradually differentiated competence in the two languages is indicative of the transition from a single undifferentiated system to two separate systems, each encompassing the knowledge of only one language.

In a radical approach to bilingualism, Swain (1977) claims that the distinction between monolingualism and bilingualism is somewhat artificial in that they are in fact not totally unrelated entities, but "realisations of a single phenomenon, in which varying aspects are observable in different degrees" (Swain, 1977:28). The underlying argument is that even monolinguals have a varied and diverse experience of their language which is made up of pragmatically different codes. Language is not a static, monolithic system but a dynamic and complex one where codes overlap and intertwine to create a sophisticated linguistic repertoire. The implication of this line of reasoning is that if the distinction monolingual/bilingual is simply an arbitrary one, then acquiring two or more languages simultaneously is not significantly different from learning only one. Swain proposes an initial stage of "mixed speech" where sentences contain elements from both languages: this stage is inevitably followed by a process of differentiation of the two languages during which the child gradually learns to separate the two linguistic systems to the point where they are completely independent of one another. In line with a principle of economy Swain argues that it is more parsimonious to assume an initial common storage of rules which will be differentiated only later on, rather than two separate codes from the start.

A real wealth of information on the acquisition of two languages from birht is Leopold's (1949) study of his daughter's bilingual acquisition of English and German. Leopold reports an initial stage before the age of two when Hildegard mixed English and German words in two- and three-word utterances, which he considers as evidence that child did not yet use "two separate systems of speech" (Leopold, 1949, vol. I). Nevertheless, Hildegard also produced fully English and fully German utterances, although quantitative data of mixed and non-mixed utterances are not presented by Leopold. Moreover, when data was collected Leopold, a bilingual speaker himself, was always present, and the German-speaking mother also spoke English although not with

Hildegard. It is therefore entirely possible that Hildegard's mixing was, at least partially, conditioned by the presence of a bilingual speaker[2].

Swain & Wesche (1975) analyse the speech of Michael, a French-English bilingual child, between the ages of 3;1 and 3;10. During the fortnightly tape-recorded sessions two researchers were always present except for one occasion when only the French-speaking investigator was present and three occasions when Michael was alone with the English-speaking investigator. During the free-play sessions Michael played with both researchers who pretended to be monolingual speakers of English and French respectively and asked him to act as a go-between and translate for them. Given the unusual situation in which the child found himelf, that of a translator for two adults, it is not surprising that Swain & Wesche found a number of instances of mixing and note that the most frequently substituted items were articles and nouns[3]. The data is divided into two chronological stages, although the authors do not specify the age range they cover. In the first period, 20% of the French substitutions into English were nouns and NPs, as for the English substitutions into French 80% of them were nouns and NPs. Interestingly, while the figure for French substitutions remained around 20% the English substitutions only decreased to 55%. Swain & Wesche point out that this discrepancy reflects "the difference in the level of development of Michael's English and French" (Swain & Wesche, 1975: 18). They also add that "[Michael's] English is 'filled in' with French functors". By this rationale we would expect a higher proportion of French articles in utterances where the matrix language is English (e.g. *Le* dog has left) as opposed to a higher proportion of English articles in otherwise French utterances (*The* chien est parti). The relevant mixed element to focus upon would therefore be the deteminer and not the noun or the noun phrase; in other words, rather than reporting the percentages of the substitutions of nouns or NPs Swain & Wesche should make clear what was the proportion of mixed determiners.

The authors also report examples of what they call "stuctural interaction", i.e. the transfer of syntactic structure from one language to the other. The interference seems to be unidirectional from French (Michael's "stronger" language) to English, as can be seen from the use of a resumptive pronoun or NP in questions:

(1a) You got a little finger, you?

---

[2] Cf. Köppe (1996: 934-35) for a similar difficulty in negotiating a strictly monolingual context when a speaker of the child' other language is present during recording sessions.

[3] Unfortunately Swain & Wesche (1975) do not report mixing percentages for the separate sessions and it is therefore impossible to know whether the absence of one of the two researchers made any significant difference with respect to mixing.

(1b) They open, the windows?

the use of postverbal negation:

(2a) I think Marcel want not to listen.
(2b) I'm going not to let you in.

possessive constructions:

(3) That's to me?

and uninverted wh-questions:

(4a) That's what?
(4b) It's what?

While, admittedly, the examples from (1) to (3) may well be the result of transfer of typically French structures, the lack of Subject-Auxiliary Inversion in (4) is also observed in monolingual children acquiring English and therefore does not need to be accounted for by transfer. It is however unfortunate that Swain & Wesche do not provide any figures for the examples above, thus there is no way of knowing whether they were sporadic or systematic occurrences.

In a study of the bilingual acquisition of four children of German-speaking fathers and non-German-speaking mothers, Redlinger and Park (1980) provide empirical results arguing in favour of the one-system approach. The authors analyse the data in relationship to linguistic development as measured by Mean Length of Utterance (MLU) (Brown, 1973), in addition to providing a distributional analysis of the mixed items in terms of grammatical category. Overall their predictions in terms of the unitary system approach are confirmed by the finding that mixing rates decrease with advancing linguistic development and with increased MLU.[4] This trend is thought to account for the transition from a single undifferentiated system to two separate systems, a process which requires the child to gain a linguistic awareness which seems to be lacking in the earlier stages of acquisition. However, in the comparison of the four children, a few discrepancies emerge which Redlinger & Park try to account for in

[4]However, as Meisel (1989) correctly points out, interpersonal comparison does not seem to corroborate Redlinger & Park's (1980) predictions. For instance at MLU 3.3, for one child, Danny, the percentage of mixed utterances is as high as 14.6 %, while for Marc it is a mere 2.1%.

terms of various linguistic and sociolinguistic factors. The high rate of mixing for two of the children, Marcus and Henrik, as compared with a third child, Marc, is ascribed to the lack of strict separation in the language input according to the one-person one-language principle where each person addresses the child in only one language and tries to negotiate as monolingual an environment as possible (Ronjat, 1913; Döpke, 1992). The authors acknowledge that the role played by the parental input could have made the difference in Marc's case where, for comparable MLUs, the mixing rates are significantly lower than for the other children.[5]

Redlinger & Park also provide a breakdown of mixed items by grammatical category where the noun emerges as the most frequent mixed item overall, followed by the adverb, article, pronoun, verb, adjective, preposition and conjunction. The authors observe that the lexical mixing encountered in the children's speech is an artifact of the delayed acquisition of certain items in one language with respect to the other. In Redlinger & Park's one-system approach, children access a single lexicon where entries are retrieved to express a given semantic concept regardless of the appropriate language feature; in other words, whichever lexical item is available to match the semantic content required will be retrieved by the children. This would suggest that children's early lexicons lack synonyms, a typical feature of a unified lexicon noticed by other researchers as well (Volterra & Taeschner, 1978; Vihman, 1985).

In Volterra & Taeschner's (1978) account of the bilingual development of two German-Italian speaking children from age 1;0 to 4;0, the unified lexicon stage is defined as stage I. Lexical acquisition seems to proceed in a similar fashion in bilingual and monolingual children, as suggested by Swain(1977). During what Volterra & Taeschner describe as stage II, two separate lexicons emerge to which a single set of syntactic rules is applied. Differentiation of the lexicon implies that children are now in a position to express the same concept in both languages; however, it is interesting to note the important role played by the pragmatic conditions in which words from the two languages are learned. The strong influence of context can be appreciated in the following two brief dialogues taken from Volterra & Taeschner (1978: 318):

Giulia (2;2)

---

[5]

CHILD	AGE	MLU	%MIXING
Marcus	2;4-2;5	2.21	21.2
Danny	2;0-2;1	2.46	12.0
Henrik	2;4-2;5	2.89	11.9
Marc	2;8-2;9	2.66	2.6

(5) Giulia: Mami, was das da? (Mummy, what that there?)
pointing to a hair pin.
Mother: Das ist eine Klammer. (This is a hair pin)
Giulia: Klammer Klammer. (and she repeats it many times)

About a month later the father, holding a hair pin, says:

(6) Father: Questa è una molletta. (This is a hair pin)
Giulia: No, non è molletta, è eine Klammer. (No, that's not a *molletta*, it is *eine Klammer*)

The cognitive task to detach one word from the specific linguistic and pragmatic context in which it has been learned is obviously a very demanding one for the child. Gradually, however, children learn to generalise and to match corresponding words in the two languages, an achievement which is demonstrated by their ability to translate between the two languages.

In Volterra & Taeschner's (1978) findings, lexical differentiation is not parallelled by syntactic separation of the two systems. Focusing on three specific areas of grammatical development (possessive constructions, adjectival constructions and negation) the authors provide evidence of syntactic mixing in the speech of Lisa and Giulia. To express the idea of possession Lisa uses the same construction in both German and Italian: e.g. *Giulia Buch* (Giulia book), *Giulia giamma* (Giulia pyjamas), *Lisa Hose* (Lisa pants), *Lisa bicicletta* (Lisa bicycle). As far as negation is concerned, only a couple of examples from German utterances are supplied by the authors, but they remark upon the Italian constructions. Lisa does produce a few sentences in which the negation precedes the verb, as in the target adult form; however, the majority of her negative constructions are of the type verb+negation as in *Fa pilli pilli no; Lisa cade no; Lisa va da la no* (Lisa 2;4-2;9). Volterra & Taeschner exclude the possibility that the Italian verb+negation constructions are the result of interference from German; if this were indeed the case we would expect *Fa no pilli pilli*, rather that the attested *Fa pilli pilli no* (Lisa 2;4). Her rule seems to be to place the negative marker *no* at the end of the sentence in both German and Italian, e.g. *Lisa machen haia haia no* (Lisa 2;7). The authors suggest that, perhaps, Lisa is simply following a pattern of development not uncommon in monolingual Italian-speaking children, where negation is placed after the verb, e.g. *piove no, sporco no* where the corresponding adult form would require the negative particle to precede the verb: *non è sporco, non piove.* In the transition to the third and final stage, children gradually realise that the two lexicons require two

distinct sets of rules; as in Swain's (1977) "common storage model", the process of differentiation consists of identifying a common rule as appropriate to a particular code thereby assigning the language-specific rules to the respective systems.

Vihman (1985) attempts an analysis of the early production of an English-Estonian bilingual child (Raivo), following the three developmental stages as described in Volterra & Taeschner (1978). Vihman's data cover a period of 21 months, from 1;1 to 2;10, during which Raivo acquired Estonian in the home and English in the community. During the first seven months (until approximately 1;8) Raivo seems to be developing a single lexicon with few translation equivalents; starting from 1;8, however, he starts to produce multi-word utterances and the gradual emergence of a dual lexicon is reported. Vihman also notes that, at the beginning of the multi-word stage, mixed utterances account for 34% of his production even though he had synonyms for 40% of his English words. More importantly, given the comparatively high level of synonymy in Raivo's vocabulary, such a rate of mixing can hardly be reconciled with Volterra & Taeschner's claim that children initially mix utterances because they still have not acquired corresponding terms in the two languages.

Over the first five months of word combination, it is primarily the use of English function words with Estonian nouns which accounts for the high proportion of mixed utterances. This finding is in accordance with Redlinger and Park's (1980) results on three bilingual children's mixed utterances[6]. More interestingly, Raivo seems to follow the pattern of the two least advanced children in the Redlinger & Park's study (Danny and Henrik), whose percentage of mixed functors is much higher when compared to the third child, Marc. On the other hand, Marc shows a higher proportion of mixed contentives in general, and of nouns in particular, a characteristic of adult language mixing too. This last observation, together with Raivo's general pattern of linguistic development leads Vihman to conclude that "language mixing by the infant bilingual prior to language differentiation is a phenomenon different in kind from what I would prefer to call code-switching by the older child and adult" (1985: 308). A possible account for the high mixing of English functors in Estonian utterances, could be provided on phonological and morphological grounds (English forms being more transparent than their Estonian equivalent), however this would merely be an *ad hoc*

6

CHILD	%FUNCTORS	%CONTENTIVES	%NOUNS
Danny	53.1	56.1	30.6
Henrik	56.1	43.9	37.0
Marc	33.3	66.6	53.3

explanation for a wider phenomenon. Vihman is more inclined to regard this recurrent developmental pattern as a reflection of the lack of linguistic awareness on the part of the children. It is not until the end of the second year of age that the child realizes adult-imposed standards of behaviour and shows his willingness to comply with them.

If, on the one hand, Vihman agrees with Volterra & Taeschner on the single-lexicon hypothesis, on the other she reports that her subject starts to separate his syntactic systems at the same time as he differentiates his vocabulary in a dual lexicon. In Vihman's analysis of Raivo's data there is no place for a transitional stage II where the lexicon is differentiated but the syntax is not. On the contrary, as soon as the child is aware of the existence of two separate linguistic entities at the lexical level, he also shows a parallel differentiation at the syntactic level.

2.2.2. The mixed-speech stage revisited

Criticism to the one-system approach has come from a host of researchers (Lindholm & Padilla, 1978; Genesee, 1989; Meisel, 1989; De Houwer, 1990, 1994, 1995; Lanza, 1992, 1997; Genesee, Nicoaldis & Paradis, 1995; Quay, 1995; Köppe, 1996; Paradis & Genesee, 1996; Genesee, Boivin & Nicoaldis, 1996; Schelleter & Sinka, 1998) on the grounds that the claim that a mixed-speech stage is evidence of an undifferentiated system is fundamentally flawed.

Genesee (1989) questions the basic assumption at the heart of the unitary system hypothesis by arguing that mixing *per se* cannot be taken to be proof of either hypothesis. In a critical review of previous accounts of mixing, he suggests context as the measure of linguistic differentiation. His argument is the following: if it were possible to establish that in early bilingual acquisition there is no differential distribution of items from the two languages as a function of context, then it could reasonably be maintained that children are not able to separate their two systems. If, conversely, it can be proved that children do in fact discriminate their use of the two languages as a function of context, then the differentiated-language systems hypothesis would gain credit. Genesee observes that in previous studies on infant bilingual acquisition not enough attention has been paid to language context. The fact that mixing decreases over time with increased MLU (see Redlinger & Park, 1980), does not necessarily imply that children move from one system to two: it could simply be that their two linguistic repertoires become increasingly more complete and solid, and they do not need to resort to borrowing any longer. In other words, mixing might be

the result of "last resort" acquisitional strategies rather than the representation of an undifferentiated system. Pursuing this line of reasoning, two alternative explanations of mixing suggest themselves: in one case the child may resort to mixing because the language being used in the context is still incomplete and does not include the lexical or syntactic resource required. In the other case the grammatical device is indeed available in the language currently in use, but the child may prefer to employ a less demanding equivalent in the other language.

In addition to the role played by specific acquisitional strategies in accounting for some instances of mixing, Genesee also invokes the all too often underestimated significance of input. Children's utterances may in fact be modelled on mixed input produced by adults or older children. The case of Bergman's (1976) daughter is reported, where the child's use of the English possessive marker *'s* in Spanish utterances (e.g. *Es de papa's*) could be traced to her nursery school teacher's use of the same construction. Other authors have acknowledged that children exposed to a high proportion of mixed input and indiscriminate use of the two languages regardless of their interlocutor, show higher mixing rates than children for whom the one person-one language principle is applied (Redlinger & Park, 1980; Lanza, 1992).

On the basis of research conducted on the perceptual abilities of infants (Jusczyk, 1982; Trehub, 1973), Genesee provides additional evidence for the ability of children to discriminate between their two languages, since differentiation "minimally requires that children be able to discriminate perceptually between the spoken languages" (1989:171). It seems in fact that infants of 6-17 weeks are able to differentiate phonetic contrasts in languages (Czech and Polish) they had never heard before (Trehub, 1973). Another study conducted on 4-day-old infants born in French-Russian speaking families (Mehler, Lambertz, Jusczyk & Amiel-Tison, 1986) has provided evidence that at this very early stage infants are already capable of discriminating between the two languages, showing a preference for French.

2.2.3. Evidence for lexical differentiation: translation equivalents

An important issue that has been neglected in those studies that have interpreted lexical mixing as an indication of lack of language differentiation, is the issue of language choice and traslation equivalents. In other words, what is the availability of lexical alternatives in the two languages in terms of translation equivalents? It may well be the case that when a child borrows a lexical item from language A to be used in

language B, she is doing so because she has no other choice, i.e. because she only has one lexical entry in either one of the two languages, but not in both. Establishing whether and from what stage of linguistic development bilingual children have access to translation equivalent is of fundamental importance in settling the dispute between one vs. two systems. A number of recent studies have addressed the issue of language choice in young bilingual children with respect to the availability of transaltion equivalents in their vocabulary.

In a study of 27 English-Spanish bilingual children in the United States, Pearson, Fernández & Oller (1995) used the MacArthur Communicative Development Inventory (CDI) to assess the children's vocabulary development between the ages of 0;8 and 2;6. All of the subjects had significant regular exposure to both languages, although input patterns varied considerably across the 27 households[7]. The main aim of the study was to test Volterra & Taeschner's (1978) claim that bilingual children go through an initial phase in which they reject cross-language synonyms. Pearson et al.'s findings show that all children but one produced doublets, the average percentage being 30.8%. Although the authors treat their results as preliminary there seems to be reason to believe that from the earliest stages of acquisition children can and will include translation equivalents in their lexicon. Pearson et al. warn however that the fact that there are translation equivalents in the children's lexicon is not conclusive evidence for two separate, independent vocabulary lists in the children's memory. They simply want to reject Volterra & Taeschner's original proposal that as long as their vocabularies are under about 75 words bilingual children avoid learning a word in a language if they already have the equivalent word in the other language.

Quay's (1995) work is along similar lines, although her methodology is very different. The data comes from a longitudinal study of a Spanish-English bilingual child, Manuela, between the ages of 1;1 to 1;10. Quay draws both on weekly videorecordings of approximately 30 minutes each for Spanish and English and daily diary records kept by Manuela's mother. By the end of the study at age 1;10 Manuela had a productive vocabulary of 300 words, 50% of which were English words (150 ENG items), 35% were Spanish words (105 SPA items), 13% were ambiguous words (38 ENG/SPA items), and 2% (6 '?' items) are words which do not seem to be either Spanish or English. In the first six months Manuela produced 9 pairs of equivalents; 36% of her English words and 41% of her Spanish words have equivalents in a period

[7] Pearson et al. (1995) do not give any information about age of first exposure to the two languages.

in which Taeschner (1983) rules out the possibility of doublets[8]. By the end of the study at age 1;10, Manuela had still the same percentage of English words with equivalents (36%), although the percentage of Spanish words with equivalents increased from 41% to 51%. Quay's sound empirical base leads her to conclude that before the age of two the vocabulary development of bilingual children is characterised by a significant percentage of equivalents, a finding in contradiction with Volterra & Taeschner's original "no equivalents" claim.

Lanvers (1997) reports similar findings in her investigation of a child being brought up bilingually in an English-German home. Data collection is similar to Quay's: diary notes collected by the mother/investigator and tape-recordings. The period studied covers 22 months from 1;1 to 2;11, the total number of equivalents ranges from 1 at 1;2 (8.5%) to 24 at 2;11 (27%) with an average of 27%. Lanvers notes that overall language dominance was found to be the most important variable in the acquisition patterns of equivalents. The child seems to be concentrating in acquiring new words in his dominant language (English) rather than learning the equivalents in his weaker language. However, when in the monolingual environment of the weaker language (German), the child concentrates on acquiring translation equivalents rather than adding new words to his German repertoire.

In a case of successive rather than simultaneous bilingual acquisition, Levy (1975) confirms the ability of children under 2 to include translation equivalents in their lexicon. Levy presents data from a longitudinal study of her son Yair from 1;11 to 2;5. Yair was brought up in a monolingual Hebrew speaking environment until his parents moved to the United States when he was 18 months, after that the child was hearing Hebrew at home from his parents and English at the day care centre and in the community at large. When Yair was 1;11 his parents/investigators started testing his ability to translate from one language to the other asking explicit questions such as "How does Pat (his day-care teacher) say X? or "How does *ima* (mother) say X? Between 1;11 and 2;5 they elicited 78 nouns, 28 verbs and 11 closed-class items. Although Levy does not state what percentage of the elicited words were supplied correctly by Yair, nevertheless she comments that "Surprisingly enough, Yair was able to perform this task appropriately" (Levy, 1975: 546).

[8] Quay's identification of equivalent terms is through the "the child's interchangeable use of one word for another to refer to the same object, event or process" (Quay, 1995: 378).

Although the data is interesting from the point of view of metalinguistic awareness and language choice in a young emerging bilingual child, nevertheless we must consider the fact that this is a case of successive bilingual acquisition where the child has had considerable exposure to only one language and is exposed to two languages simultaneously only after 1;6.

2.3. Pragmatic differentiation in bilingual children

Recent empirical studies focusing on sociolinguistic aspects of language mixing, and on its implications for the issue of language separation, have shown that bilingual children as young as two can discriminate between monolingual and bilingual speakers, and that they are sensitive to the requirements of their linguistic environment.

Lanza (1992, 1997) provides detailed evidence for young bilingual children's subtle pragmatic competence with respect to language choice and language mixing. In a longitudinal study of a Norwegian-English child (Siri) aged between 2;0 and 2;7, Lanza focuses on two aspects of the child's mixed utterances: the formal aspect (grammatical vs lexical mixing), and the context in which mixing occurs.

In accordance with Vihman's (1985) predictions, four-fifths of Siri's single item mixes are functors. What is peculiar to Siri is that grammatical mixing is unidirectional: only Norwegian grammatical morphemes co-occur with both languages, while English grammatical morphemes are restricted to English lexical utterances. Lanza takes this pattern to be an indication of Siri's dominance in Norwegian. This is also reminiscent of Genesee's (1989) prediction that "if the differentiated language-systems hypothesis were true, one would expect to find more frequent use of items from the weaker language in contexts where that language is being used than in contexts where the stronger language is being used, even though items from the stronger language might predominate in both contexts" (1989: 166). At the time when Siri inserted Norwegian functors into English lexical utterances in many cases she had already acquired the English equivalent, and therefore mixing cannot merely be the result of an incomplete single system where items from the two languages are in complementary distribution.

As far as lexical mixing is concerned, the logical expectation would be to find the same dominance of Norwegian in English utterances; however, contrary to this prediction, Siri's lexical mixing is higher with her Norwegian-speaking father than with her English-speaking mother. Lanza accounts for this unexpected situation by

resorting to the dynamics of the negotiation of the language context. Even though Siri's parents claim that they are both committed to keeping a one person-one language approach when interacting with their child, different patterns emerge in a careful analysis of the discourse context. Lanza identifies five basic parental discourse strategies (minimal grasp, expressed guess, adult repetition, move-on strategy and code switching)[9], towards child language mixing which may be placed on a continuum ranging from a strictly monolingual context to the acceptance of code switching between the two languages. The underlying argument is that children as young as two are already extremely sensitive to repair cues provided by their interlocutor, and if adequately prompted they are able to locate the trouble spot in the conversation and make the required adjustments. In her conversations with Siri, her mother is always very careful to negotiate a strictly monolingual context by pretending not to understand the child's mixed Norwegian words and asking for clarification. Her father, on the other hand, is more likely to accept the use of English words in a Norwegian context, and the conversation will carry on without requests for clarification even though he will often provide the Norwegian equivalent in his reply.

Lindholm & Padilla (1978) have found similar effects of the negotiation of context between bilingual children and bilingual experimenters who pretended to be strictly monolingual. The authors have observed that one of the reasons for the children's switching from one language to the other was to verify whether the "monolingual" experimenter was really monolingual. In sum, the available evidence suggests that while Siri's interactions with her mother are essentially monolingual contexts, her father is more willing to negotiate a bilingual context where code switching is acceptable. This interpretation would therefore account for the unexpected lexical mixing pattern in a Norwegian context.[10]

De Houwer's (1990) analysis of the language use of the young bilingual English-Dutch child in her study also reveals a strong tendency for the child to use Dutch with the people who address her in Dutch (the investigator and her father), and to use English with her English-speaking mother. The addressee is a major determinant in the selection of the language to use for Kate, although the topic, but only in the case of colour terms, is also a factor. A large number of English words in otherwise Dutch

---

[9]See Lanza (1992) for a more articulated description of the five discourse strategies.

[10]Lanza's (1992) perceptive analysis as evidence of a differentiated language system in bilingual children is however biased by the age of the child under consideration. Before Lanza, Vihman (1985) had observed that two-year-olds have already gained a sufficient self-awareness and metalinguistic awareness to be credited with two separate systems. The crucial question seems rather to be whether, prior to this stage of cognitive development, children's representations are in fact separate or whether a qualitative change comes about around the end of the second year of life.

utterances are in fact colour terms that she has learnt mainly at her English-speaking nursery school. What is even more revealing of the sophisticated pragmatic knowledge of this young bilingual child is that not only is she sensitive to the language different interlocutors actually use with her, but she also appreciates varying degrees of bilingualism in the people around her. Overall there are more mixed utterances that have English words inserted in Dutch utterances than the other way round; moreover while more than half of the 18 Dutch utterances directed at her English-speaking mother are simply single words, the picture is more complex in the case of the English utterances addressed at the bilingual investigator and at her father, the language is much more complex and varied. De Houwer explains this pattern by appealing to the different degree of bilingualism of Kate's interlocutors. While the investigator and the child's father are fluent Dutch-English bilinguals, her mother's Dutch is very limited and she does not use it very often. Kate must obviously be aware of her interlocutors' proficiency, or lack thereof, in the two languages, and although she normally replies in the language she is addressed in, she also knows how far she can go in using the other language, depending on whom she is talking to.

In the analysis of three French-German bilingual children from the DUFDE project, Köppe (1996: 948) reports that "[f]rom very early on onward, all three children show that they know in which language they are expected to interact".[11] Köppe adds that the children often hesitate before mixing, and sometimes they explain that they do not know the word in the appropriate target language:

(7) Ivar, 3;01.03

G: was is das für ein tier?
"what kind of animal is this?"

I: ein, eine TORTUE
"a, a TORTOISE"

G: ja und wie sag ich dazu? das ist richtig aber wie sag ich dazu?
"yes and what do I call it? that's right but what do I call it?"

[11] The acronym DUFDE stands for Deutsch und Französisch Doppelter Erstsprachwerb (German and French - Simultaneous Acquisition of Two First Languages). The project investigates the simultaneous acquisition of German and French in a number of preschool children growing up in middle class families in Hamburg, Germany. The children were videotaped every second week from the beginning of linguistic development (1;0-1;6) to the age of at least 5 years. French is the native language of the mothers, while the fathers' native language is German. Each parent used his or her respective native tongue when communicating with their children.

I: i weiß nich mehr.
"I don't know anymore"

There is a degree of individual variation in the extent to which the three children in Köppe's study resort to mixed utterances, and the author argues for an input-driven explanation. Although all the parents in her study reportedly use a one-person one-language strategy, they vary to the extent in which they negotiate a more strictly monolingual environment, as opposed to a more flexible bilingual environment where mixing is tolerated. Children are very sensitive to direct or indirect information about the acceptability of mixing and to the bilingual vs. monolingual competence of their interlocutors, and they behave accordingly.

Although a number of studies have shown that children as young as two can differentiate their language use as a function of their interlocutor (De Houwer, 1990; Lanza, 1992, 1997; Meisel, 1994; Köppe, 1996; Genesee, Nicoladis & Paradis, 1995), Genesee, Boivin & Nicoladis (1996) argue that the children's pragmatic behaviour could be explained by their familiarity with their interlocutors, either parents or well-known investigators. In the attempt to disentagle the familiarity variable from true communicative competence, Genesee et al. (1996) recorded four English-French bilingual children (average age 2;2) in interaction with unknown monolingual strangers in their weaker language; in addition the children were also separately recorded with their bilingual parents. For two of the children the unfamiliar interlocutor was a monolingual French speaker, and for the other two it was a monolingual English speaker. The results of the study confirm previous research findings that young bilingual children can use their languages differentially and appropriately with familiar interlocutors (i.e. their parents). Moreover, the analysis of the children's behaviour with the strangers also confirms that young bilinguals can also accomodate their language use with interlocutors with whom they have had no previous contact, thus showing an even higher degree of sensitivity to the pragmatics of bilingual discourse.

In sum, there is a substantial body of evidence suggesting that bilingual children can differentiate the use of their two languages in context-sensitive ways from the age of two onwards. Nonetheless, most bilingual children do use mixed utterances in their speech, this behaviour has been explained both in connection with the child's own linguistic resources, i.e. the availability of translation equivalents, and in relation to the kind of monolingual vs. bilingual context negotiated with the adult interlocutor.

The next section reviews some of the studies that have advocated separate morphosyntactic development in Bilingual First Language Acquisition.

2.4. Separate morphosyntactic development

Complementing the findings from a more sociolinguistic perspective, several recent studies investigating the early grammatical development of bilingual children have given increasing credit to the idea that formal language differentiation takes place very early in BFLA.

The notion that grammatical development in a bilingual child's two languages proceeds in a separate and independent fashion has been proposed by De Houwer (1990: 339) in the Separate Development Hypothesis:

> [T]he morphosyntactic development of a pre-school child regularly exposed to two languages from birth which are presented in a separate manner proceeds in a separate fashion for both languages.

Her detailed analysis of Kate's morphosyntactic development in Dutch and English, together with a comparison with monolingual peers, leads the author to conclude that the two languages do develop separately and independently. There is no evidence that the child's morphosyntactic knowledge in one language functions as the basis for the emergence of a grammatical system in her other language, or viceversa. De Houwer (1990, 1994, 1995) also makes a number of crucial methodological points for the test of the Separate Development Hypothesis. Firstly, only syntactic phenomena that differ significantly in the two languages should be investigated. Secondly, only lexically and morphologically unilingual utterances can form the empirical basis of such an investigation; and thirdly, the interpretation of developmental stages must take into account any available data from monolingual acquisition. Only those developmental "errors" that are not the same in normal monolingual acquisition can be taken as evidence for crosslinguistic influence.

Convincing evidence for a differentiated system in child bilingual acquisition also comes from Meisel's (1989) study of the syntactic development of two French-German speaking children, (P and C). Meisel focuses on word order, a linguistic phenomenon which is sufficiently distinct in French and German to serve as a test case to decide for or against the one-system hypothesis.

German is canonically described as a verb-second (V2) language in that the finite verb obligatorily surfaces in the second position of a main clause as in (1):

(8) Ich schreib einen Brief.
"I write a letter."
*Schreib ich einen Brief.
"Write I a letter."
*Ich eine Brief schreib.
"I a letter write."

Placement of the verb in second position is also required whenever some other element (NP, Adverb, wh-word, complementizer) is topicalized:

(9) Heute schreib ich einen Brief.
"Today write I a letter."

(10) Einen Brief schreib ich heute.
"A letter write I today."

Topicalization of objects and adverbs is possible in French as well, and the resulting sequences show the verb in third position: AdvSV(O) or OSV(Adv). In German, on the other hand, topicalization of adverbs and objects results respectively in AdvVS and OVS sequences, given the verb-second requirement. Meisel's data show that both children correctly leave the verb in third position in French and place it in second position in German, thus demonstrating a clear-cut distinction between their two syntactic systems.

With respect to the acquisition of word order, Köppe (1996) reports the emergence of language-specific VP word orders in two French-German children. Moreover, not only did the children produce VO utterances in French and OV in German, but whenever they mixed elements from both languages the order of the verb and the object appears to be detemined by the choice of language for the verb[12]:

(11) e COQUILLAGE reinlegen (Pascal, 1;10.0)
"a SHELL inside-put"

[12] Note however that Ivar, the third child in Köppe's study, produces VO and OV word orders in each of his two languages. Unfortunately Köppe does not provide additional information on Ivar's concurrent use of VO and OV structures and it is therefore impossible to assess the importance of the phenomenon. She does however say that the lack of a fixed language-specific ordering inside the VP is not sufficient to claim that Ivar is operating with only one system, as he seems sociolinguistically and pragmatically aware of dealing with two languages.

(8) on regarde DAS (Pascal, 1;8.22)
"we look at THIS"

(Köppe, 1996: 942)

Additional evidence for language-specific development is offered by Sinka & Schelleter (1998) who present data on the morphosyntactic development of a German-English bilingual child (Sonja, 1;11-3;8), and an English-Latvian bilingual child (Maija, 1;2-3;0). For both children they show a lead-lag pattern whereby morphosyntactic marking in the form of verbal inflections, plural inflections and case marking emerges later in English than either German or Latvian. Both children acquire language-specific inflections in both languages, and no crosslinguistic influences are observed in the children's languages that develop as separate and independent systems. The authors emphasise the typological differences of the input in English, German and Latvian, and ascribe the later emergence of morphosyntactic marking in English for both children to the relative morphological poverty of the English input an to its paradigmatic inconsistency. In German and in Latvian, on the other hand, the consistency and the transparency of morphological marking and of distributional evidence facilitates the children's grammatical development.

Shifting from the verbal to the nominal domain, Koehn (1994) analyses French and German noun phrases in Ivar, one of the French-German children also included in Köppe's study. The linguistic input plays an essential role in Ivar's production in his two languages. Focussing on those elements which provide the most accessible and reliable information, the child employs different mechanisms to express the grammatical feature [number]. In French he begins by marking number on articles, while in German he concentrates on noun suffixes. Note that, apart from nouns ending in *-al* and *-ail* whose plural suffix is *-aux*, in French the plural is formed by adding /s/ to the singular noun, a phonetically invisible process[13]. Therefore, from a perceptual point of view, number is clearly marked only on articles in French, and it is precisely on articles that Ivar first marks number distinctions. By contrast ,in German articles are not unifunctional, also because of the additional Case feature. For example, *die* is the nominative and accusative form of the definite feminine singular article and also the nominative and accusative form of the plural article. Given the ambiguity of the German article paradigm it is only natural that the child should start by focussing on a

[13] There are of couse exceptions, such as a number of suppletive plural forms: e.g. *oeil/yeux, monsieur/messieurs, etc.*, in addition to exceptions, e.g. *chandail/chandails, rail/rails, détail/détails,* etc.

more reliable paradigm such as the one provided by noun suffixation for plural formation[14].

A comparison of the nominal and verbal domain of two French-English bilingual children, Yann and Mathieu (Paradis & Genesee, 1996a) reveals significant differences only in the verbal domain. The emergence of verbal inflections is attested around MLU 1.50 in French and only later in English ("at the end of First Syntax" (Paradis & Genesee, 1996a: 26) which following Valian (1992) they define as the developmental stage where MLU is between 1.10 to 2.00)[15]. As for the nominal domain there are no dramatic differences between the children's English and their French. In both languages the children use definite and indefinite determiners, percentages in obligatory contexts are comparable across the two languages, and although their repertoire is still very limited, Paradis & Genesee credit them with productive use of determiners at First Syntax.

These findings, the authors claim, provide a strong argument for a weak continuity approach to language acquisition whereby the instantiation of functional categories is not a prerequisite of acquisition, nor is it subject to maturational constraints, but it is driven by the learning of lexical items and the associated functional properties. Similar findings are replicated by another study on French-English bilingual acquisition (Paradis & Genesee, 1996b). Investigation of finiteness, negation and the use of pronominal subjects in three French-English bilingual children between 1;11 and 3;3 confirms the language-specific nature of bilingual acquisition. In all three children the researchers found a large gap in the appearances of finite utterances between French and English, with English lagging behind. This fact combined with the absence of English utterances with postverbal negatives of the type "I eat not" lead the authors to conclude that the children are approaching the two languages as two separate and self-contained problem spaces. The children do not transfer their emerging knowledge of subject-verb agreement and verb movement from French to English, in essence they proceed as if they were acquiring the two languages as monolingual children. However, although the vast majority of studies in the grammatical development in bilingual first language acquisition support the Separate Development

[14] See Koehn (1994) and Köpcke (1988) for the proposal of a possible continuum of schemas for plural nouns on the basis of cue strength determined by perceptual salience, type frequency and reliability.

[15] It is not entirely clear from the presentation of the data when exactly the Inflectional Phrase (IP) would emerge in the children's English. For Yann there are no finite verbs by the end of the study (age 3;0), for Mathieu finite verbs are attested around 20% by the end of the period of observation at age 2;11.

Hypothesis, some researchers have suggested that a weaker version of the SDH would capture the data more realistically[16].

2.5. The role of crosslinguistic influences in BFLA

At the present time a large proportion of researchers in BFLA share the belief that bilingual children develop separate grammatical systems from the earliest stages of syntactic development. Nonetheless, some have suggested that the notion of two separate grammatical systems is not incompatible with the possibility of contact between the two languages (see Müller, 1998). The hypothesis that young bilingual children acquire two separate systems, does not automatically imply that these two systems are necessarily impermeable to one another, some degree of contact can potentially be envisaged. The idea is still somewhat controversial in the field, and it is in need of a more precise definition. There is much debate, for example, as to whether such contact is systematic, or whether it is a characteristic of individual children. A recent article by Müller (1998) investigating the acquisition of word order in German subordinate clauses reports that only some of the bilingual children studied, with German as one of their two languages, make errors in verb placement. Some do not. Moreover, the type of errors are not unique to bilingual children, but are qualitatively similar to those made by some monolingual German children, although Müller claims that they are more frequent in bilingual children. Müller's argument is that bilingual, and monolingual, children make errors in the face of ambiguous input, in this case the V2, non-V2 asymmetry between matrix and subordinate clauses. She argues that the developmental error German-speaking children make is to locate the [finiteness] in the Inflectional Projection, rather than in the higher Complementiser Projection, as in adult German. Although the assumption underlying this error is qualitatively the same in bilingual and monolingual children, in bilingual children who are also exposed to languages such as Italian, English, and French, where indeed [finiteness] is associated with the Inflectional Projection, this wrong analysis is reinforced by evidence coming from the other language, hence the larger number of errors in bilinguals.

Döpke (1997a, 1997b, 1998) also argues for crosslinguistic influences in the face of ambiguous input in a study of three German-English bilingual children in Australia. The data is presented following stages of development defined by MLU rather than by

[16] De Houwer (1995) goes as far as stating: "To my knowledge, there are no methodologically sound studies of morphosyntactic development in young bilingual children that have found any evidence against the separate development hypothesis" (De Houwer, 1995: 240).

chronological age (MLU ranging from 1.75 to 5.75+). Between MLUs 1.75 and 2.74 the emergence of word order with respect to verb placement in initial or final position and negation develops separately for the two languages and according to the monolingual norm. However, in a subsequent phase, between MLUs 2.75 and 4.74, considerable variation with respect to monolingual German children starts to appear. While by this stage monolingual German children clearly differentiate between non-finite verbs in sentence final position and finite verb in non-final position (Clahsen & Penke, 1992; Clahsen, Penke & Parodi, 1993/94), the bilingual children in Döpke's study moved verbs to prenegative position irrespective of finiteness, something unattested in German monolingual acquisition. Döpke's claim is that while monolingual German children only have to distinguish between finite verbs in non-final position and non-finite verbs in final position, and they do so once they have acquired the relevant morphosyntactic correlates of finiteness, bilingual German-English children also have to contend with the distinction between non-final verbs in German and non-final verbs in English. The English input reinforces the 'verb before object' cue for German and generates cue conflict as to where the non-final verb is structurally located: V or C. The children finally resolve the puzzle when they start paying attention to the relative order of verbs and negation in the two languages.

Recently, Hulk & van der Linden (1997, 1998) have presented data from Anouk, a French-Dutch bilingual child being raised in a one-person, one-language environment and they have also argued for crosslinguistic influences, although more indirect ones than those proposed by Müller and Döpke. French and Dutch have different word orders, French is an SVO language while Dutch is OVS, hence, Hulk & van der Linden argue, if bilingual children start out with separate syntactic systems we should only observe V XP utterances in French and XP V utterances in Dutch. The data show a different picture: from 2;3.13 to 2;7.5 V XP and XP V word orders are equally distributed in Anouk's French, the deviant XP V order decreases in the period between 2;7.5 and 3;6.25 and is unattested after 3;6.25. The authors suggest that while the earlier instances of XP V order in French can be accounted for by the absence of a higher Inflectional Projection, the later ones, also marginally attested in monolingual French acquisition (Ferdinand, 1996), can be viewed as instances of topicalisation[17]. Although this topicalisation mechanism is equally operating in the monolingual acquisition of French, Anouk resorts to it with a much higher frequence than her

[17] Cf. Deuchar (1997) for a similar account of the early two-word utterances of a Spanish-English bilingual child, and Köppe (1996) for the observation that "With regard to Ivar, however, language-specific combination of verbs and objects cannot be attested before the emergence of INFL" (Köppe, 1996: 943).

monolingual peers. Why should this be the case? Hulk & van der Linden suggest that around age 3 there is an asynchrony in Anouk's syntactic milestones in the two languages, while in Dutch left-peripheral elements are analysed as filling the specifier of the Complementiser Projection (CP), in French they are still adjoined as there is no clear presence of a CP projection and subject clitics are still optional. Given this disparity between the two languages the argument is that the more advanced language, in this case Dutch, "bootstraps" the acquisition of a similar structure in the weaker language. Gawlitzek-Maiwald & Tracy (1996) make a similar claim for Hannah, a bilingual German-English child, and her use of mixed functional elements such as German modals and auxiliaries in her English utterances until the relevant functional elements are acquired in English around age 2;9[18].

2.6. Summary

The issues raised by the notion of crosslinguistic influencences in BFLA are important ones for a better understanding not only of *what* children acquire, but also of *how* they acquire it. The notion that a bilingual child is the sum of two monolinguals offers a somewhat limited view of how knowledge may be organised and accessed by individual speakers. At the present time there seems to be reasonably convincing evidence that bilingual children do approach the language learning enterprise by assuming that languages that are perceptually, morphologically, syntactically different, and which are presented to them in a separate fashion according to the one-person one-language strategy, are indeed different entities whose rules and idiosyncracies will have to be learnt in a language-specific way. Nevertheless there is also reason to believe that there must be a degree of osmosis between the knowledge that is acquired for one language, and the knowledge that is gained for the other. The challenge for BFLA researchers is to define more precisely what makes such crosslinguistic influences possible, whether there is a degree of individual variation, whether there are morphosyntactic devices that are more likely than others to be transferred, and what exactly is the nature of the ambiguity.

[18] Paradis & Genesee (1996) have specifically investigated the issue of "bilingual bootstrapping" in their longitudinal study of two French-English bilingual children between 2 and 3 years of age. They claim that "The large gap between french and English in the use of finite utterances and the absence of English postverbal negatives indicate that the children are not transferring the verb movement parameter from French into their English grammar, nor is the presence of French accelarating their acquisition of English syntax" (Paradis & Genesee, 1996: 19).

The next chapter explores some of the issues regarding the upbringing of bilingual children with specific reference to Carlo (C.), the informant of this case study. A detailed description of C.'s linguistic environment is presented together with information on the methodology used for the collection, the transcription and the coding of the data.

Chapter 3

A case study in bilingual first language acquisition: input patterns and methodology

3.1. Methodological premises

The empirical data presented in this thesis come from a single English-Italian bilingual child. The decision to use only one child for this study was motivated by a number of factors that I will now illustrate and justify.

The nature of the research question, i.e. the investigation of the emergence of functional categories in a bilingual environment, imposed restrictions on the type of informants and data required. This study's focus on the emergence of early syntax made it necessary to collect data longitudinally from children younger than two.

Another issue to address in the adoption of the most appropriate methodology for the data collection, was the choice between an experimental approach vs. a naturalistic approach. The decision to use naturalistic data/spontaneous speech was motivated by the age of the children suitable for the study. The intention was to capture the critical transition from the emergence of speech at the one-word stage, to the initial phases of the multi-word stage, in Brown's (1973) terminology from Stage I to Stage III. Although there is considerable individual variation, in the second year of life most children will start producing one-word utterances followed by two and multiword utterances around their second birthday. Hence the necessity to find children younger than two in order to record crucial information about the earliest stages of the emergence of grammatical relations. When dealing with such young children, an experimental approach may not always be appropriate. Experimental conditions have often been shown to be far too demanding for two-year-old children, and tests at this age are not so much tests of natural behaviour, but of test behaviour (Donaldson, 1978; Clark, 1982).

A further requirement of this study was that the children were being brought up in a *bilingual environment,* and specifically with English and Italian as their two languages. Therefore the search was further narrowed down to children that would satisfy this additional criterion .

3.2. Bilingual environments

The aim of this study is to research the emergence of syntax in two languages, Italian and English, where morphological and syntactic correlates differ along a number of dimensions. In Italian, Person deixis is expressed by a rich verbal agreement system, and subjects may be null, when overt they may occupy the preverbal or the postverbal position, according to different discourse and pragmatic requirements. Number and gender are marked on articles, nouns, and modifiers; bare nouns are only allowed in lexically-governed positions, and they receive a default existential interpretation. In English, the verbal agreement system only marks Number and not Person, Person deixis is therefore expressed by overt subjects, null subjects are only allowed as instances of topic drop in root position. Grammatical gender is only marked on 3 p.s. personal pronouns, and number is overtly marked only on nouns. Bare nouns are allowed with plural and mass generics. Of course, these are only some of the morphosyntactic differences between Italian and English, but they will suffice for the purposes of the present study.

Differently from other crosslinguistic studies comparing and contrasting monolingual children speaking different languages, the approach taken here is to consider cases of bilingual first language acquisition (BFLA) as defined by De Houwer (1990: 3):

> BFLA refers to those situations in which
>
> (a) a child is first exposed to language B no later than a week after he or she was first exposed to language A, and
> (b) a child's exposure to languages A and B is fairly regular, i.e. the child is addressed in both languages almost every day.

There are clearly various input patterns that qualify a child's environment as bilingual (see Döpke, 1992: 12-13). It could be the case that both parents speak the minority language at home between them and with the child, and the child is exposed to the language of the community through nursery, playgroup and other activities outside the family home. In this case the minority language would be the home language, where a predominantly monolingual environment would be negotiated in the family,

and the language of the community would be spoken mostly outside the home environment.

Alternatively, it could be that one parent is a native speaker of the minority language and speaks to the child in his or her mother tongue, and that the other parent is a speaker of the language of the community and speaks to him in that language (one-parent one-language strategy); in this case the home environment would be more of a bilingual environment than in the case in which both parents speak the minority language[1]. Of course there are other input sources other than parents that are equally important and need to be taken into account when assessing a child's bilingual input, such as nursery caregivers, childminders, friends and extended family. The two scenarios just sketched above obviously do not exhaust the many and varied possible bilingual environments a child may be brought up in; other factors may complicate the picture. For example the presence of other siblings who may or may not be bilingual; extended visits from monolingual relatives; whether the parent who speaks the language of the community in the one-parent one-language scenario also understands and speaks the minority language spoken by the other parent. In addition to the number and diversity of sources that are available to the child growing up in a bilingual environment, another issue of crucial importance is the relative access the child actually has to these different sources of input. In the case in which both parents are speakers of the minority language and negotiate a monolingual environment in the home, one might expect that the consistent and exclusive use of the minority language as the home language will constitute robust enough input for the child to start using that language with her family while keeping it separate from her other language that she will mainly speak in the community.

In the one-parent one-language approach, when only one of the parents is a speaker of the minority language, this process of identification and language separation may prove more difficult and at times unsuccessful. The fact that one of the parents speaks to the child in a language different from the language of the community does not automatically imply that the child will grow up speaking that language, while he will certainly grow up speaking the language of the community he lives, studies and works in. Establishing a successfull bilingual environment requires a conscious and deliberate effort on the part of the parents. If only one of the parents is a native speaker of the minority language a number of strategies may be adopted in order to maximise the child's exposure to this language and hence increase the likelihood that she will acquire

[1] For a definition of the one-parent one-language strategy see Ronjat (1919) and Döpke (1992).

it at the same time as the language she is exposed to in the community. The presence of other caregivers who also speak the minority language such as childminders, relatives, friends etc., is an excellent opportunity both to increase the quantity of the child's input and also to expose her to a variety of speakers so that the minority language will not only be "mummy's" or "daddy's" language, but the language of an extended community.

Yet another crucial factor is whether the parent who speaks the language of the community also understands and speaks the minority language spoken by the other parent. When this is not the case then it is likely that the minority-language-speaking parent will switch to the community language when the other parent is present so as not to exclude him or her from the conversation. This results in minimal input in the minority language to the child, restricted to the times when she is alone with the minority-language-speaking parent, or possibly other speakers of that language, but not in the company of other monolingual speakers who have no understanding of it. Of course it could also be that the parents have decided to carry on with their one-parent one-language strategy regardless of the presence of other monolingual interlocutors. This strategy has been known to work as long as children are still very young and the kind of verbal interactions parents have with them are still quite limited in scope (Döpke, 1992). It is not difficult to imagine a French-speaking mother talking to her 12-month-old baby son in French in the presence of other monolingual English-speaking people and then code-switching to English to talk to them. On the other hand, it is a less likely scenario that the same mother will address her 5-year-old in French in the presence of his monolingual English-speaking father thus partially excluding her partner from their conversation. Moreover, as children grow older interactions become more complex and articulated; the number of participants to a conversation increases to involve more than just one interlocutor at a time. In the case of bilingual children this means that the language combinations may vary dramatically to include bilingual speakers, monolingual speakers of the minority language and monolingual speakers of the community language. Even parents who have always strictly adhered to the one-parent one-language strategy report difficulties when children become older, typically when they start attending school, and their social network includes an increasingly large number of monolingual speakers of the community language. In other words, adhering to the one-parent one-language strategy (at least in the home environment) requires a certain amount of cooperation on the part of the parent speaking the language of the community. Ideally she or he would at least understand the minority language if not speak it; this would allow the other parent to use his or her language

knowing that this would not jeopardise effective communication between all the family members.

3.3. Bilingual environments defined

When I started looking for suitable children for this study I contacted various associations and institutions in Edinburgh that could provide me with access to bilingual families where at least one of the parents was an Italian native speaker. This first approach proved unsuccessful. By placing an advert in the Edinburgh University Bulletin, a publication circulated among staff and students of the university, and by word of mouth, I finally contacted three families where one of the parents (in two cases the mother and in one the father) was a native speaker of Italian and where the children (two girls and one boy) were at the time younger than two and were at the one-word stage.

The parents were asked to fill in a questionnaire on the use of the two languages in the family in order to assess to what extent the children had access to both languages and whether they used them both[2]. Unfortunately in all three cases the children did not qualify as actively bilingual in that, although they seemed to understand and respond appropriately to Italian most of the time, they invariably used English and only occasionally imitated a few Italian words. Common to all three families was the fact that the English-speaking parent understood little or no Italian and spoke only English to the child and to the other parent. In the boy's case (T.), the Italian-speaking mother switched to English when the father was present, while in the two girls' cases (A. and E.) the Italian-speaking parents reported continuing to use Italian with the child while switching to English when addressing the other parent. In T.'s case the strategy adopted by the mother had the effect to reduce the Italian input to those occasions when mother and child were together alone. This, combined with the fact that T. attended an English-speaking nursery every morning for four hours and that he spent three afternoons a week with an English-speaking childminder, clearly created an imbalance in the amount of Italian input the child had access to.

As for the two girls, the fact that the Italian parent continued to address the child in Italian regardless of the presence of the English-speaking parent might lead one to expect a more balanced and successful bilingual environment. However this was not the case for A. and E.. It is likely that the ability, and the willingness, of the parent

[2] A copy of the questionnaire is reproduced in Appendix 1.

speaking the language of the community to speak, or at least understand, the minority language spoken by the other parent will have an impact on the child's perception and affective evaluation of the minority language. Language is learnt in a social environment and because of this it is associated with a number of communicative, cultural and emotional values that one cannot abstract from when studying the mechanisms of its acquisition. The inability, or the unwillingness, of one parent to speak or at least understand the other parent's language may constitute an impediment in establishing the minority language as equally worthy and necessary in the eyes of the child. Moreover, the minority language parent's bilingualism makes it even less compelling for the child to learn a language that she only hears in very restricted contexts. A parent who is obliged to switch to the language of the community when the other parent is present deliberately relegates his minority language to a subordinate position, an imbalance that is most likely perceived by the child. Failure to convey to the child the importance and the worthiness of the minority language will almost certainly contribute to make the acquisition process a difficult task.

Another variable which must not be underestimated is the actual *access* the child has to a balanced bilingual input. In the case of the three children above, they all attended an English-speaking nursery school for at least three hours a day. T. was looked after by an English-speaking childminder three afternoons a week and spent the rest of the time either alone with his Italian-speaking mother or with both parents when only English was spoken. A. spent the morning at home with her Italian-speaking mother and three hours at an English-speaking nursery in the afternoon. The rest of the time she was cared for by a bilingual English-Japanese childminder who used only English with her. In the evenings and at week-ends she spent time with both parents hearing Italian from her mother and English from her father. E. also started to attend an English-speaking nursery at 2, previously she was looked after by her mother at home. Overall E. spent an estimated 70% of her waking time in a monolingual English-speaking environment.

In all three cases, aside from occasional contact with Italian grandparents once or twice a year, the Italian-speaking parent was the only consistent source of Italian input. No substantial provisions were made for Italian-speaking childminders or other alternative ways to increase the children's everyday exposure to the minority language. The implications of this strategy are twofold: firstly, the input the Italian parent alone can provide is not sufficient, considering that the overwhelming majority of linguistic

input to the child is in English, through the other parent, nursery caregivers and other childminders.

There seems to be some tentative evidence for a lower limit of exposure below which language learning, or at least language use, seems to be threatened. In a recent study of lexical learning in bilingual infants, Pearson, Fernández, Lewedeg and Oller (1997) report that even children who have a very limited exposure to one of their two languages (for some as little as 20% or less of their waking time) will still learn vocabulary items in proportion to the amount of exposure to that language. They concede, however, that "whether children can acquire a grammar and a sound system from low levels of exposure to a language cannot be inferred from vocabulary learning and will have to be specifically addressed in future studies" (Pearson et al., 1997: 55). The children in the Pearson et al.'s study who heard one of the languages 20% or less of their time were generally very reluctant to use that language and seemed to be "tuning it out" in the play sessions. Pearson et al.'s conclusion is that there does seem to be a threshold at around 20% exposure below which children were not at ease with that language and did not use it willingly or spontaneously. If this were the case, the reduced access that T., A. and E. had to the Italian input could be one of the factors involved in their failure to acquire and use the language.

Secondly, if the Italian parent is the only identifiable source of Italian input in a predominantly English-speaking environment, it will be more difficult for the child to grant Italian the same status of language of a community rather then being simply "the way mummy or daddy speaks". On this subject I think an anecdote reported by the parent of a bilingual child will be very revealing. The child, a successfull bilingual, was the first born of an Italian mother and an American father living in Scotland. When he was almost three the family took a trip to Italy to visit the maternal grandparents in Rome. One day, accompanied by his mother, he went to the local playground where he met other children and started playing with them, when he suddenly realized that they too spoke Italian. In a state of great excitement he ran back to his mother exclaiming: "Mum, here all the children speak Italian!" This discovery obviously amazed the child since, although in Scotland he had regular contact with a number of Italian-speaking adults, there were no other children that spoke Italian. With no previous evidence to the contrary the child had probably come to think that Italian was a language spoken by a selected minority of adults. Going to Italy it suddenly dawned on him that the communtiy speaking Italian was in fact much larger than the circle of parents, adult friends of the family and childminders that he knew in Scotland. Creating a *realistic*

linguistic environment is therefore crucial to the development of the child's linguistic identity in the minority language, a demanding task that requires a considerable committment on the part of the parents.

It is reasonable to conclude that due to restricted access to linguistic input and the lack of positive reinforcement of Italian as the language of a real community, the three children T., A. and E. were not acquiring Italian as well as English, even though they had been exposed to both languages from birth.

After excluding T., A. and E. I had to continue my search for other actively bilingual children younger than two. The ideal candidates had to be younger than two, already at the one-word stage, and had to be using both English and Italian to qualify for the study. Eventually the child selected for the case study was 17 months of age at the beginning of this study, and at the time he had not started producing any meaningful utterances in either language. Given these premises there was an element of risk involved in that the child could have fallen in the same pattern as T., A. and E. and not acquire any Italian. As will be explained below, however, the bilingual environment the child was being brought up in was offering fairly substantial indications that the child would indeed acquire Italian as well as English.

3.4. Carlo, the Target_Child

I had known Carlo (henceforth C.), the child that was to become the informant of this case study, since my arrival in Edinburgh. Initially I had ruled out his participation to the study because of his age - at the time he was only one year old and I required children who were closer to their second birthday. Because the child had not started speaking yet, there was no guarantee that he would in fact become bilingual. As in the case of T., A. and E. he could have grown up in a bilingual environment without necessarily becoming and active bilingual. C. however had an older brother, M., who was five years old at the time and had become a very successful and balanced bilingual. There was reason to believe that the younger brother would follow in his footsteps and become as successful a bilingual. C.'s parents felt very strongly about bilingualism and they had decided they would try in every possible way to raise their children in a positive and stimulating bilingual environment. In the same way as they had done with M., C.'s older brother, they both spoke to C. in Italian (although the father is American he speaks fluent Italian). In addition, since both of them worked full-time they enlisted the help of Italian childminders to look after the children when they were out of school

and nursery. They clearly had a very positive attitude towards the importance of bilingualism which they did their utmost to convey to their children. When I approached C.'s parents about his potential involvement in the study, they expressed a keen interest in the research project and offered their support enthusiastically.

3.5. C.'s exposure patterns to English and Italian

3.5.1. C.'s linguistic background

C. is a second-born child and the youngest in a family consisting of father, mother and two brothers. Although C. is a second-born child he has in fact two siblings: brother M., four years his senior, and step-brother A., ten years his senior. C. was born in Edinburgh, Scotland, where he lives with his family. C.'s mother is an Italian native speaker and a very fluent speaker of English. C.'s father is a native speaker of American English and a fluent speaker of Italian. M. is an English-Italian bilingual speaker, and A. is a monolingual English speaker with some basic knowledge of Italian.

In order to maximise exposure to Italian, English being the language of the community, both of C.'s parents decided to speak Italian with him, M. also speaks Italian with C., while A. speaks both with him.[3] When not addressing C. the parents normally speak English between them, the father will also speak English with A. and M., and the mother will speak Italian to M. and English to A.. Besides adopting Italian as the only language both parents use with C., they also decided to enrol him at a local nursery school on a part-time basis so that the child would only spend the morning with English-speaking caregivers and the rest of the day with an Italian-speaking childminder until the parents come home around 6 p.m.[4]

In addition to his immediate family there are a number of people with whom C. has come into regular contact in his first two years of life. His maternal grandparents normally spend the month of September visiting in Edinburgh; they are strictly monolingual Italian speakers and therefore exposure to Italian increases when they are present since conversations that C.'s parents would normally have in English between them are in Italian for the benefit of the grandparents; moreover both grandparents can

[3] C.'s father spoke mostly English to him until he was approximately 5 months old, then switched to Italian.

[4] C. started attending a creche when he was 5 months old; subsequently he remained in the same nursery until he was 3;2, and he then moved to another nursery nearby.

only address C. in Italian. C. also visits his grandparents in Rome twice a year, each time for a period of 2-3 weeks. Overall then exposure to Italian significantly increases during the 8-9 weeks of the year C. spends with his grandparents either in Rome, or during their visits to Scotland.

Other Italian speakers who have been in close contact with C. over the last four years include a number of Italian babysitters who alternate in taking care of C. and of his brother M. in the afternoons between 1, when C. gets picked up from nursery, to 6 when his parents come home. The Italian babysitters speak exclusively Italian with the children, occasionally alternating between English and Italian if some of M.'s English-speaking friends are invited to the house.

To sum up, in his daily routine C. has access to several adult bilingual speakers (parents, babysitters, Italian-speaking friends of the family now resident in Scotland), to a child bilingual speaker (his brother M.), to a number of receptive child bilingual speakers, to a number of adult monolingual English speakers (caregivers at the nursery, family friends) and child monolingual English speakers (children at the nursery, M.'s friends) but to no child monolingual Italian speakers. C. does not have Italian cousins for instance, nor does he have any Italian friends belonging to his age group.

3.5.2. C.'s week-day routine

C.'s getting up routine involves moving from his bedroom to his parents' where he enjoyes playing with his mother and brother M. for about half an hour, while his father makes breakfast. When the children are ready the whole family sits down for breakfast, which involves both English and Italian being spoken. The radio is usually tuned to Radio 3 (a classical music station, hence little speech), and occasionally Rai Uno (an Italian TV channel) is on. After breakfast C. is taken to the nursery by one of his parents.

C. spends four hours a day at a local nursery between 9 a.m. and 1 p.m., except for a period of eight months from when he was 1;4 to 2;0, when he also spent one afternoon a week at the nursery. At the nursery C. is in a stable group of peers which varies from 6 to 9 children, 1 to 2 caregivers are normally in charge.

At 1 o'clock one of the Italian-speaking babysitters in charge picks C. up from nursery where he has already had his lunch. A. and M. are not normally at home at

lunch time therefore mostly Italian is spoken with some English between C.'s parents (when they come home for lunch). In the afternoon C. spends a couple of hours with one of the Italian childminders before they pick up M. from school; normal activities include looking at books, playing with jigsaw puzzles, dominos, wooden cubes, listening to music, drawing. Between 3 and 4 C. and the childminder walk to M.'s primary school where they wait with other English-speaking mothers and children. Occasionally C. plays with the other children while waiting for M., but hardly ever speaks to them. When the children get home they normally have a little snack and then, weather and time of the year permitting, they may go to the nearby park for an hour or so. C. does not normally speak to other children in the playground but is sometimes willing to play with them, and even though he does not speak to them he obviously overhears conversations going on around him.

Occasionally M. has English-speaking school-friends over to play with him which means that more English is spoken. However, before M. is picked up from school, C. spends two to three hours (depending on whether M. stays on for "after-school club") in the company of his childminder which means a considerable amount of one-to-one interaction and of speech addressed exclusively to C.. For two to three hours a day the childminder's attention is entirely devoted to C. and to his needs; unlike at the nursery, there are no distractions coming from other children or demands such as preparing food, answering the telephone, talking to colleagues, tidying up, etc. Even when M. comes home the ratio adult-child is still 1-to-2 and there is the added benefit of Italian input coming from M. to C. and between M. and the Italian-speaking childminder.

Most of the activities in which C. engages with his childminders have some focus on language: looking at books and naming objects, animals and activities; drawing and naming the objects drawn; playing with wooden cubes and "explaining" what is being built; listening to music and singing along. During the time C. spends with his childminders he is constantly provided with linguistic stimuli tailored to his particular interests and needs.

When C.'s parents come home around 6 p.m. there are no organised activities as such, the children are usually in the kitchen while the parents are busy cooking. When the food is ready the whole family sit down for dinner which means that both English and Italian are spoken. C. is addressed in Italian by both parents, by M. and occasionally by A., while the rest of the family switch between English and Italian.

After dinner C. may go off to another room with M. and A. or with a visitor, which normally means more English than Italian. Bathtime for C. and M. is around 7.30-8 p.m.; C.'s mother is in charge, hence only Italian is spoken. No bed-time stories are read to the children either in English or Italian, C. normally falls asleep around 8-8.30 p.m..

3.5.3. Family life at week-ends and on holiday

Routine at the week-end is slightly different from the rest of the week. Both parents are at home from work and can look after the children. On Saturday mornings, after breakfast, the children usually watch some children's programmes on Italian satellite television, then go along with their father to the local supermarket for the weekly shopping trip. In the afternoon they may all decide to go for a walk nearby or into town, sometimes an Italian-speaking childminder comes for a few hours to look after C. and M.. On Sunday mornings a group of Italian-speaking friends come to visit with their children, none of whom actually speak much Italian. Visitors occasionally come for a meal at the week-end, on the whole English-speaking guests are more frequent, but Italian-speaking ones do pay visits too.

Because both parents are available at the week-end, the linguistic environment is decidedly bilingual as opposed to a more rigid subdivision into a strictly monolingual English context (e.g. nursery in the morning during the week) and a strictly monolingual Italian context (e.g. Italian-speaking babysitters in the afternoon). Both parents, M. and A. use both languages when the family are all present, while C. is exclusively addressed in Italian.[5]

The family daily routine also changes during family holidays to Italy or the United States, and visits from the maternal grandparents who are strictly monolingual Italian speakers. Over the year the family usually spend their Christmas holidays in Rome visiting C.'s maternal grandparents where the exposure to Italian is obviously increased through immersion in the community.

During the family's annual summer holidays to the United States and Canada, the single most significant difference from the situation in Scotland is the absence of Italian childminders and of a strictly monolingual context for five hours a day. The

[5] A. uses English predominantly however.

parents are always present and therefore the Italian input is maintained through them and through M., but the general pattern is of a more bilingual environment with an overall increased exposure to English. The presence of English-speaking cousins is also an important socialising factor. The time C. would normally spend playing with an Italian-speaking childminder is by and large replaced by games with his English-speaking cousins with the obvious consequence that the amount of English input is significantly greater than at home in Scotland. The effect of this increased exposure to English was particularly noticeable when C. came back from his holidays in the United Stated in the summer when he was 2 and a half years old. For the first time since the beginning of the recording sessions in English, he showed a greater willingness to speak English, to initiate conversations and to volounteer information without being constantly prompted by the adult interlocutor.

3.5.4. Features of the Italian input

The picture emerging from the information above is one where relative exposure to Italian and English changes over the year, with peaks of Italian input during trips to Italy and visits from maternal grandparents, and peaks of English input during the summer holidays in North America. The constant remains the regular, balanced exposure to the two languages throughout the year with a variety of bilingual and monolingual speakers who help C. establish a sense of community and identity. Although C.'s bilingual environment appears well balanced with a fair access to both languages on a daily basis, the quality of the input is in fact somewhat different between English and Italian. Italian is largely the home language spoken by his family and childminders with whom he has a close one-to-one relationship. English is the language he hears and speaks at the nursery where he shares the caregivers' attention with at least another five or six children.

The kind of attention C. receives from his parents and his childminders who are there exclusively for him and for his brothers, must surely be different from the kind of attention the nursery staff can devote to him. It goes without saying that the emotional bond between parents and their children is an exclusive and privileged one which has no equal. Linguistic communication creates and strengthens this bond throughout the child's early years in a unique way thus attributing to language an affective as well as a communicative value. In a similar fashion, a close one-to-one relationship with one or more childminders can develop into a significant emotional attachment over time. Although this cannot be on the same level as the child's attachment to his parents, it

constitutes a important part of the child's social and affective world and as such contributes to the child's developing sense of identity and community membership. Thanks to one-to-one everyday contact, C. has had the opportunity to develop close relationships with most of his Italian childminders, they have become an important part of his everyday social and emotional life. For at least 2 to 3 hours a day C. has the undivided attention of one of his Italian caregivers, during which time he is constantly exposed to language through games, book readings, videos and various other activities ranging form preparing a snack together to going for a walk in the park or to the shops. C. is continually engaged in interaction with an Italian-speaking adult, and he is her constant focus of attention.

The situation in a nursery environment cannot replicate the close one-to-one relationship C. has with his family and his Italian babysitters. In the transition from a dyadic to a polyadic situation, the nature of the interaction between adult and child necessarily changes. The following two sections present an analysis of the features of the particular kind of linguistic input children receive in group situations. The implications of differential exposure patterns for children growing up bilingually are discussed at the end of section 3.6.2.

3.6. Input patterns in polyadic conditions

3.6.1. Adult verbal behaviour in one-to-many situations

The previous section has provided information about the pattern of C.'s exposure to the two languages, with particular reference to the kind of interaction he engages in in Italian, the home language. In this section we will look more in detail at the nature of the input he receives in English through daily nursery attendance. Exploring the differences and the similarities between the home environment where C. is exposed predominantly to Italian, and the nursery environment, where he is exclusively addressed in English, will contribute crucial information for the assessment of the child's use of his two languages and his grammatical development in both.

A large number of naturalistic and experimental studies on the interaction between input and language acquisition have shown that joint involvement between the child and her adult interlocutor is a crucial prerequisite for lexical learning (Tomasello & Todd, 1983; Tomasello & Kruger, 1992; Tomasello, Strosberg & Akhtar, 1996; Ambalu, Chiat & Pring, 1997). In other words, it is essential that children be able to coordinate attention to the

caregiver with attention to the object or event of interest in order for learning to take place. Moreover, in addition to shared attention, the quantitiy of language heard (Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991), and the usefulness of that language as a source of information about the grammar (Snow, 1989; Pine, 1990) are two other crucial variables. Caregiver's maintenance of semantic contingency, and use of a conversation-eliciting style are key factors in the establishment of an adult-child discourse strategy which provides the child with the ideal means for extracting information on language structure and language use. A recent study on shared attention and grammatical development in a group of 30 normally developing children (1;2-2;7), and a group of 6 autisitc children (3;6-7;9) (Rollins & Snow, 1998: 671), showed that, at least for the typical children, maternal conversational style contributes to the development of the child's grammar and "children whose mothers engage in more child-centred, less directive talk progressed more rapidly toward grammar".[6]

If joint involvement and child-centred, conversation-eliciting parental style are so important for early lexical acquisition, and subsequent grammatical development, as observed in a number of studies on parent-child dyadic interaction, what of those situations where the adult-child ratio is not 1:1, but one adult must divide his or her attention between a number of children? The nursery environment is a typical setting where one adult normally looks after a number of children and therefore cannot devote his or her entire attention to only one child, but must share it between several children.

A number of studies have focused on the difference between dyadic and polyadic situations, and the ways in which adult-child interaction varies not only in quantitative terms, but also in qualitative terms when the adult-child ratio is one to many (Tizard, Phillips & Plewis, 1976; Schaffer & Crook, 1979; Sylva, Roy & Painter, 1980; Tizard, Carmichael, Hughes & Pinkerton, 1980; Schaffer & Liddel, 1984; McDonald & Pien, 1982; McCartney, 1984; Pellegrino & Scopesi, 1990). The main findings common to these studies of polyadic adult-child interaction are the following: firstly, as one would expect, the quantity of talk specifically addressed to individual children diminishes drastically in polyadic situations, although the amount of talk overall increases. For instance, Sylva et al. (1980) report that in their study each child was exposed to an average of three minutes of talk per hour. Secondly, in polyadic situations, nursery staff show a tendency to treat the children in their group as a series

[6] Unlike for the typical children, for the autistic children, maternal conversational style did not contribute significantly to the children's grammatical development. Rollins & Snow (1998: 670) comment that "[t]his may well reflect the more general social deficits of children with autism, leading to their lack of attention to social cues and difficulties in maintaining attention to social stimuli".

of dyads, rather than as a homogeneous group, and the number of group-addressed utterances is typically smaller than individual-addressed utterances. Thirdly, the adult's responsiveness to children's bids for attention decreases in polyadic situations. Schaffer & Liddell (1984) report that, while in the dyadic situation the great majority of children's bids were acknolwedged, in the polyadic situation, four out of every ten bids were ignored. Fourthly, the adult's conversational style changes substantially from dyadic to polyadic situations. In polyadic situations there is a tendency on the part of the caregiver to use more prohibitive and peremptory language than in dyadic situations, and at the same time there is a smaller number of conversation-eliciting moves, and the bigger the group the smaller the number of semantically contingent expansions and open questions (Pellegrino & Scopesi, 1990).

All the children in the studies mentioned so far are children who are growing up in a monolingual environment, and therefore it is highly likely that any deficiencies in the reduced amount of attention and talk each child receives in the group situation will be largely compensated by interaction with other caregivers outside the nursery setting. In the case of a bilingual child such as C., for whom the nursery is the only environment where he has constant and exclusive access to one of his two languages, the effects of quantity and quality of input in a polyadic situation are likely to be more visible. In the following section we will present analyses of data collected in two different nursery settings, when C. was aged 2;4 and when he was 3;1. Although the results here are in line with previous research on child-directed talk in a group care situation, a breakdown of the results by the individual children in the group also reveals interesting differences that are obscured by group analysis. A breakdown by individual children reveals a significant unevenness in the distribution of child-directed conversational behaviour. Implications for C.'s linguistic development will be discussed in light of these findings.

3.6.2. Child-directed talk in two nursery settings

In order to establish what kind of linguistic input C. received in English outside the family home, three hour-long video recordings were made at the two local nurseries the child attended during the data collection period. The focus of the observation was the nursery's staff behaviour with the children in their everyday activities, their responsiveness to the children's bids for attention, their conversational style, and whether they addressed the group as a whole, or whether they tended to single out

individual children. In order to minimise disruption, and to avoid making the caregivers self-conscious about their behaviour, an unmanned videocamera was positioned in one corner of the nursery room, and staff were told that the purpose of the videorecording was to collect data about how C. interacted with his peers outside the home.

One one-hour videorecording was made at the first nursery C. attended when he was 2;4, and two one-hour recordings were made at the second nursery he attended. Two samples of five minutes each were randomly selected from the first recording twenty minutes apart, and four five-minute samples were randomly selected from the second set of recordings. In the first nursery (nursery one), one adult is in charge of 8 children, while in the second nursery there is one staff and 6 children (nursery two). The samples were transcribed in CHAT format and every utterance was coded with a dependent tier %add indicating the addressee of the utterance. Each adult utterance was then classified into one of four categories following Pellegrino & Scopesi 's (1990) coding system for the functional analysis of adult utterances:

(a) EMPATHETIC BEHAVIOUR: utterances by means of which the adult expresses approval and support for the verbal and/or non-verbal behaviour of the child (repetition, offer of availability, confirmation, consolation, verbalisation of activity suggested by the child)

(b) CONVERSATIONAL BEHAVIOUR: utterances specifically aimed at eliciting or maintaining verbal interaction with the child (utterances with phatic function, e.g. oh, ah-ah, etc., open questions, comments on objects or events.

(c) ORGANIZATIONAL BEHAVIOUR: utterances designed to orientate and/or modify the activity and/or attention of the child (commands, requests for attention, prohibitions)

(d) DIDACTIC BEHAVIOUR: utterances whose aim is to provide the child with new information, and/or verify the child's knowledge (descriptions, closed questions, requests for repetition)

Figure 1 and Figure 2 below show the division of all of the adult utterances into the four communicative behaviours just described:

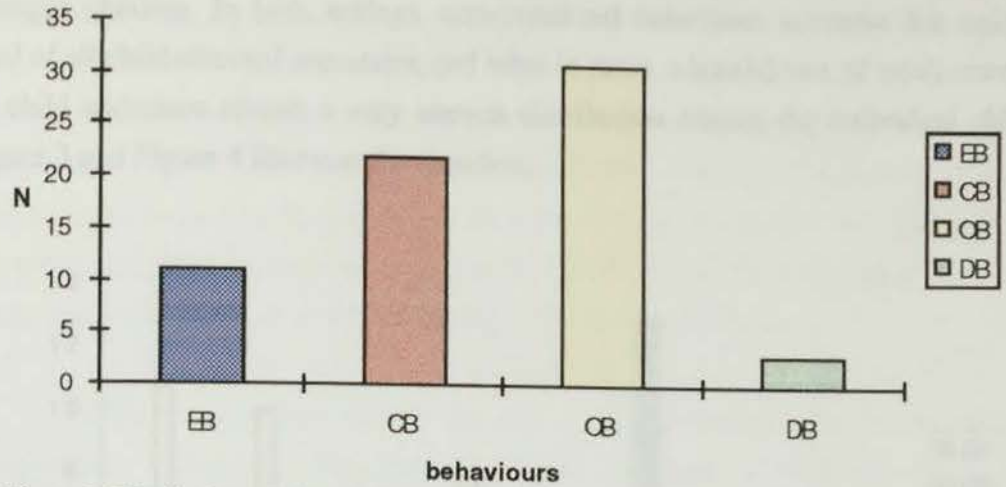


Figure 1. Distribution of child-directed utterances in the nursery one setting (EB = Empathetic Behaviour, CB = Conversational Behaviour, OB = Organisational Behaviour, DB Didactic Behaviour)

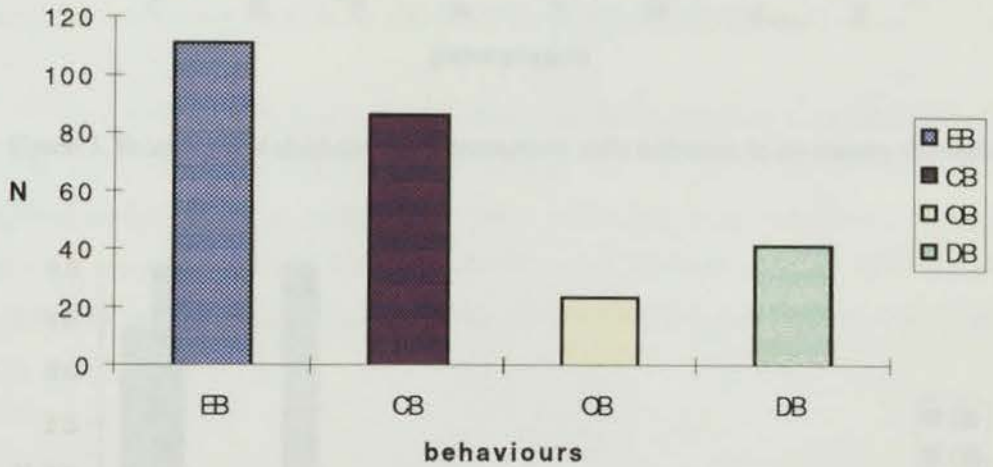


Figure 2. Distribution of child-directed utterances in the nursery two setting (EB = Empathetic Behaviour, CB = Conversational Behaviour, OB = Organisational Behaviour, DB Didactic Behaviour)

In both nursery settings conversational utterances account for the second largest group, 32.83% in nursery one, and 32.95% in nursery two. However, while in nursery one the largest proportion of utterances are organisational utterances (46.26), in nursery two the largest percentage of utterances are empathetic ones (42.04%), and organisational utterances account for a mere 9.09%. The discrepancy in the proportion of organisational utterances between the two settings is likely to be biased by the ages of the children in the two groups. While the average age in the nursery one setting is

around 2, in the second setting children are aged 3 and older, hence there is less need to control and restrain their behaviour to the same extent as that may be necessary with younger children. In both settings, conversational behaviour accounts for only one third of all child-directed utterances, and what is more a breakdown of adult utterances by child addressee reveals a very uneven distribution among the individual children. Figure 3 and Figure 4 illustrate the situation.

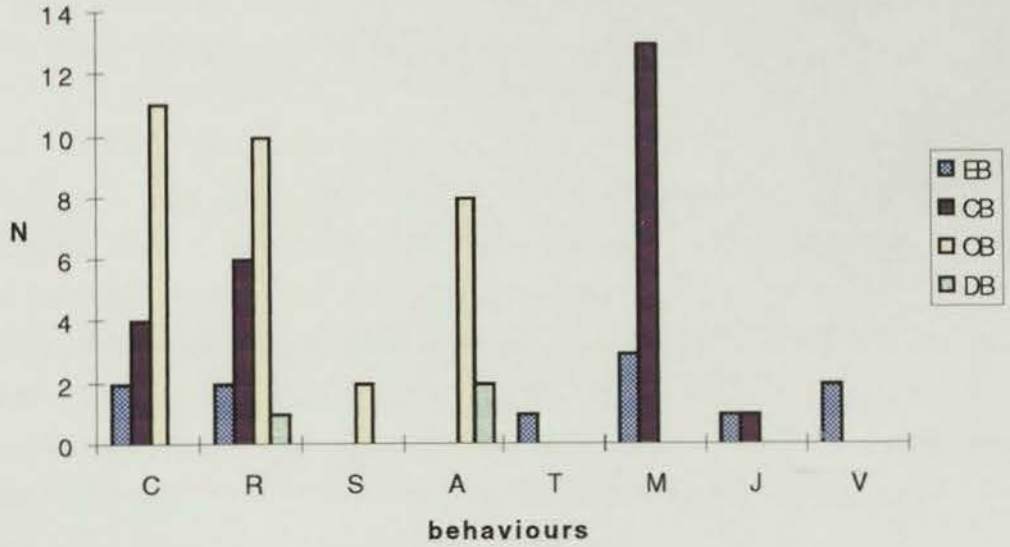


Figure 3. Breakdown of child-directed utterances by child addressee in the nursery one setting

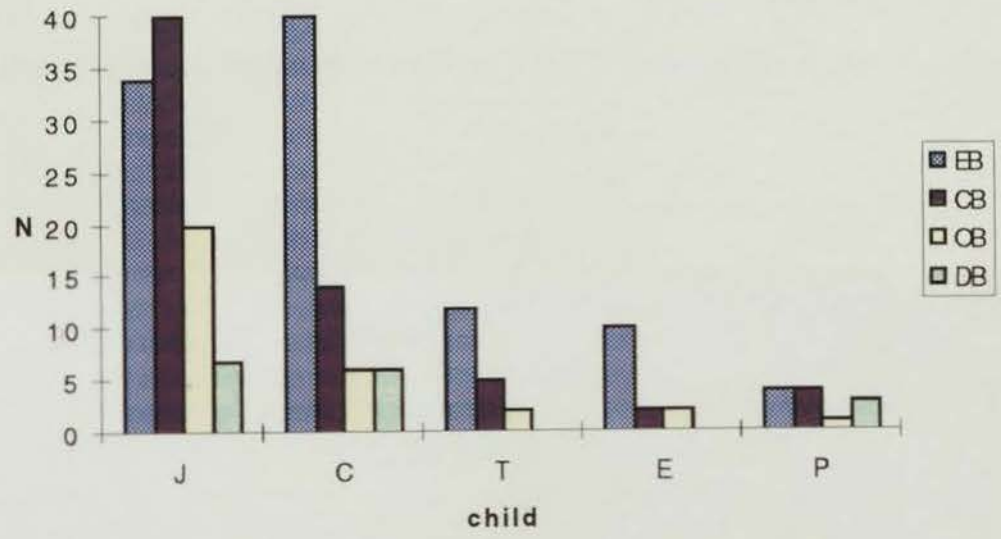


Figure 4. Breakdown of child-directed utterances by child addressee in the nursery two setting

In Figure 3 and Figure 4, the charts plot the number of all adult child-directed utterances by individual child and by category. In the nursery one scenario, C. is one of the four children who are on the receiving end of a number of conversation-eliciting utterances, he is the recipient of 18,18% of such conversational moves, while one child alone, M., is the addressee of almost 60% of all conversational behaviour, and another child, R., is the target of another 27% of conversational utterances. The large proportion of conversational utterances addressed to M. is likely to be a function of the child's age, she is the oldest in the group and the one who is more linguistically advanced, while the other children are still at the one-word stage, or at the beginning of the multiword stage.

In the nursery two setting we find a similar scenario as far as conversational behaviour is concerned. One child, J., is the addressee of 46% of all conversational utterances, while C. is the direct target of a conversational utterance only 16% of the time, and the children as a group are addressed conversationally 23% of the time. Both charts clearly show a certain unevenness in the proportion of different types of behaviour children are the recipients of, in particular of the crucial child-centred conversational utterances, such as open questions, and semantically contingent expansions, which are thought to be the verbal behaviour that is most conducive to optimal child learning.

In the attempt to find out what could be at the basis of such individual differences in the distribution of child-directed speech, the children's behaviour and the adults' responsiveness to it was examined in more detail. Figure 5 and Figure 6 report for each child the number of child bids for attention, the number of adult bids which are not a response to a child bid, i.e. the number of adult initiatives, and the number of ignored child bids.

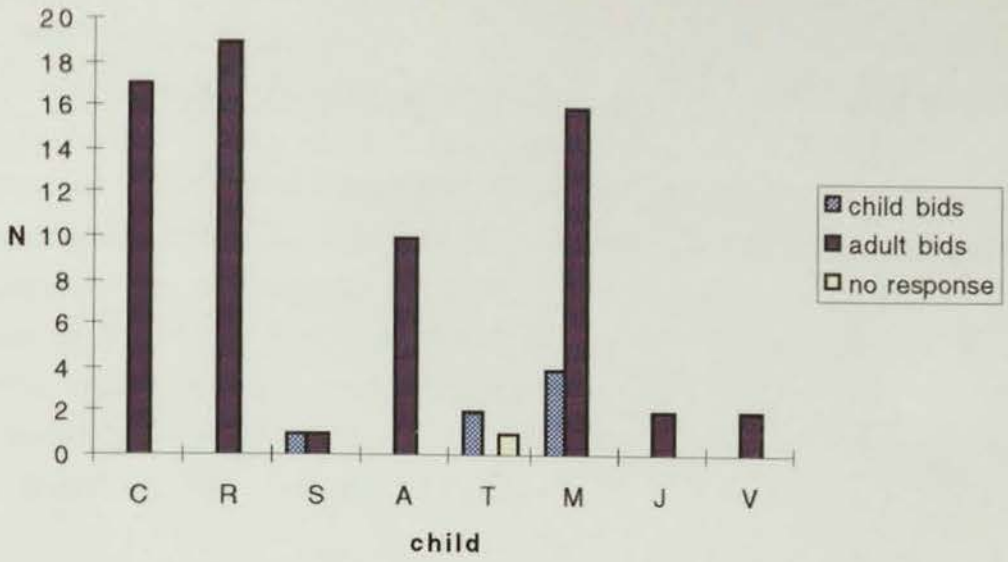


Figure 5. Number of child bids, adult bids for each child, and unresponded child bids in the nursery one setting

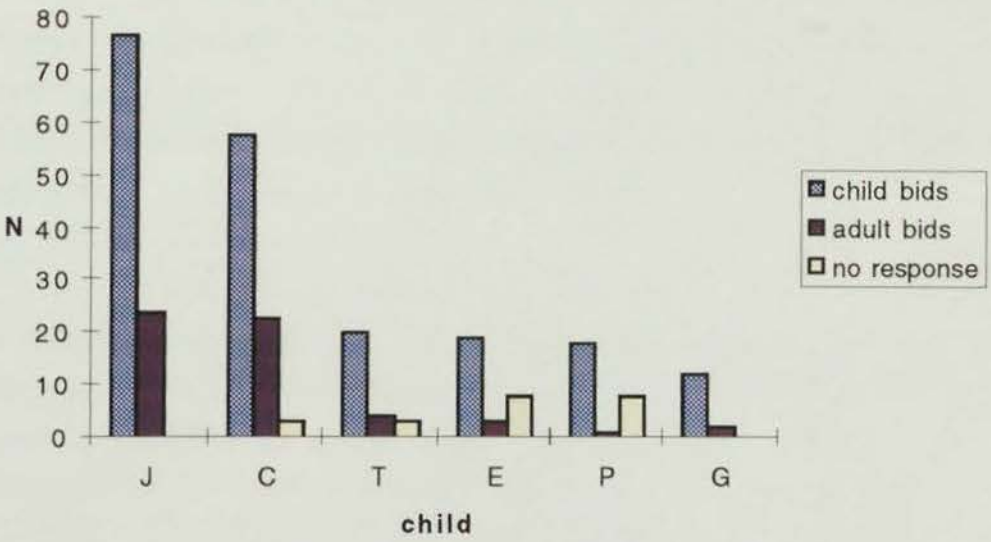


Figure 6. Number of child bids, adult bids for each child, and unresponded child bids in the nursery two setting

In Figure 5, the nursery one setting, M., the child that was addressed the largest proportion of conversational utterances, is also the child with the highest number of bids. Five out of the eight children never make any bids for attention, while C. has the second largest proportion of adult bids (25%). Of the three children that do make bids for adult attention, only one is ignored on one occasion. In Figure 6, the nursery two setting, J., the child who receives the largest proportion of conversational utterances is

also the one who makes the largest number of bids for the adult's attention. C. makes the second largest number of bids, and C. and J. together account for more than 60% of all child bids, while the other children make a substantial smaller number of bids, in the region of 10% each. J. and C. are also the two children that receive the largest proportion of adult bids, but while for J. the majority of these adult bids are conversational utterances, for C. they are empathetic utterances, such as praise and repetition of the child's utterance. C.'s bids, however, are only ignored 5% of the time, while for three other children, T., E., and P., the proportion of ignored bids is much higher, around 13% for T., and 30% for E. and P. Moreover, not only are these children's bids ignored more often, they are also the ones that receive a very small number of adult bids compared to J. and C.

The picture that emerges from this analysis of child-directed input in a nursery environment is one where conversation-eliciting behaviour accounts for a mere 30% of all speech addressed to children. Moreover, not all children are treated equally, some receive proportionally more conversational stimulation than others. M. in the nursery one setting, and J. in the nursery two setting, receive a number of conversational utterances that is by far greater when compared to that received by C. and the other children in the group. Children who make fewer bids for the adult's attention are also the ones that are more likely to be ignored when they do, and the ones who are less likely to be on the receiving end of the adult's initiative.[7]

In line with previous studies, it was found that nursery caregivers tend to treat the group as a series of dyads. In the nursery one setting there are no utterances addressed to the group as a whole, and in the nursery two setting only 14.39% of utterances are addressed to the whole group, the rest are addressed to individual children.

In sum, this analysis of C.'s nursery environment has shown a situation in which the child's access to stimulating input, rich in semantically contingent expansions and recasts, is somewhat limited in this polyadic situation. The English input is certainly not of the same quantity and quality as the Italian input the child receives at

[7] In their analysis of age biases in polyadic situations, Pellegrino & Scopesi (1990: 105) also note that "the adult takes up, often with expansions and corrections, the verbal bids of those children whose language is more developed". M. and J. are indeed the most advanced children in group 1 ad group 2 respectively, and their bids are often in the form of open questions, or attempts to engage the adult in conversation. The three children in the nursery two setting who tend to be ignored more often, are less advanced linguistically, their bids are often monotonous repetitions of the same statement, e.g. *Look at this!* that do not require the adult to give a full answer.

home from his family and his Italian-speaking babysitters. This discrepancy in the access to stimulating, child-centred input between the two languages must obviously be taken into consideration in any explanatory account of the child's grammatical development in the two languages. The analysis and the discussion of the emergence of formal grammatical features in C.'s English and Italian in chapter 6, 7, 8, and 9, will factor in the differential exposure to the languages as a non-trivial source of information and potential explanation of the different course of development in the two languages.

The next sections will present information on the data collection, transcription and coding of the linguistic material.

3.7. Data collection

3.7.1. Methodological choices

When I first considered C.'s participation to the study he was 17 months and was still in the prelinguistic phase where he vocalized but did not produce any meaningful linguistic utterances as such. Although he had not started using language, he was in the age range at which children normally have either started producing their first words or are about to. Therefore, with the parents' consent, it was agreed that I would start visiting the house once a week, normally on a Sunday morning, so that I could start collecting data for the research project.

I decided to use a videocamera for the recording sessions for three principal reasons: firstly, in the case of such a young child there was a good chance that some or most of his early utterances would be difficult to understand without the appropriate pragmatic context, and a video support did prove invaluable when it came to transcribing the data[8]. Secondly, nonverbal behaviour is important to assess the child's communicative competence in participating in a conversation. At times non-verbal behaviour can be an alternative rather than an accompaniment to verbal behaviour, and as such an indicator of the child's competence and willingness to use the language (Bates, Camaioni & Volterra, 1979; Ochs, 1979). Finally I anticipated the need to view

[8] For all the recordings I used a Sony Handicam video Hi8, mostly in combination with a tripod, occasionally balanced on a chair or on a table depending on the location of the videorecorded session.

the tapes at different intervals subsequent to recording when the transcripts and possible field notes would not have been as informative.

It could be objected that the use of video equipment is more intrusive than audio equipment, and could conceivably influence the behaviour of those who are knowingly being videotaped. The presence of a videocamera never seemed to bother C.; at the beginning of the study he was too small to understand the significance of the object. Later, as time passed, he started to realise that the camera had something to do with images and occasionally he would ask to see himself on camera but by and large he seemed oblivious to the presence of the object during the recording sessions. Moreover, being familiar with the child in other situations that did not involve the use of a videocamera, I can say that his behaviour was no different in those situations and that being videotaped did not make any noticeable difference in his behaviour, linguistic or otherwise.

The first four recordings were spaced over a period of two months and took place on Sunday mornings at C.'s house. C., his mother, myself and occasionally his brother M. would be present. The videocamera was normally balanced on a shelf and C.'s mother and myself would try and elicit some speech by showing the child his favourite toys and picture books. His mother would perform a routine by asking him questions such as *Dov' è il maialino?, "Where is the piglet?",* while holding a book with a picture of a piglet within easy reach or a toy animal, and C. would point to the appropriate picture or toy animal. His comprehension of nouns as measured by the correct identification through pointing to pictures or touching objects was excellent, however the only sound he ever produced in these four initial sessions was /*ka*/ , some form of undifferentiated deictic he used when pointing to pictures or toy animals. The recording sessions were discontinued for 2 and a half months and resumed when the family came back from their summer holidays in North America. By that time it was agreed that I would become one of C.'s childminders for one afternoon a week (1 p.m. to 6 p.m.), which proved an ideal opportunity to become a familiar figure in his home environment and also gave me the chance to form a more comprehensive picture of his personality and his development.

During the summer C. had started to say his first words, therefore systematic recordings could now start. The plan was to record him twice a week for approximately 30-40 minutes, once with an Italian-speaking interlocutor and once with an English-speaking interlocutor. Because I was going to spend one afternoon a week with the

child, I did the recordings in Italian myself, although occasionally other Italian childminders also took part to the recordings. For the English data, Karen, a Canadian friend of C.'s family who knew the child well, kindly agreed to assist me. While the recordings in Italian at C.'s house, either with myself or another Italian childminder, constituted a representative sample of C.'s daily interaction in Italian, the recordings at home with an English-speaking adult were a forced choice dictated by some practical considerations. C.'s daily *direct* exposure to English was through attendance at a local nursery school. Therefore, a more representative sample of his daily contact with English speakers would have been a series of recordings at the nursery where he interacted with staff and children. After two initial attempts I had to discontinue recordings at the nursery as the quality of the material I collected was extremely poor. At the time C. was in a group with six or seven other children younger than two supervised by two staff in a fairly small room. Given the children's age and the lack of collaborative play, the verbal interaction between them was minimal. Virtually all verbal exchanges were between staff and children, and when the nursery caregivers spoke to the children individually, most of their interventions were of a directive or prohibitive nature, rather than aimed at initiating and/or sustaining a conversation. The opportunity to record any speech from C. was therefore minimal in this situation. Also there were frequent outbursts of crying and screaming from younger children which made most of the material virtually impossible to transcribe with any degree of accuracy. The alternative was then to have an English-speaking adult come to the house to play with C. and record these free-play sessions.

3.7.2. Organization of the recording sessions: the participants

The four tables below provide an overview of the 39 recording sessions which constitute the Carlo corpus. Table I and Table II list the file numbers, together with the age of the child, and the Mean Length of Utterance in words (MLUw).[9]

[9] The MLU is computed in words rather than in morphemes (Brown, 1973) for a more balanced comparison between the two languages. Computing the MLU in morphemes would bias the results in favour of Italian where words tend to be polymorphemic. For example the Italian word *camminiamo, "we walk"*, is constituted by two morphemes, the stem *cammin-* and the inflectional 1 p.p. ending *-iamo*, while the English equivalent *walk*, is only one morpheme.

file	age	MLUw
1	1;10.8	1.071
2	1;10.27	1.165
3	1;11.17	1.360
4	1;11.25	1.444
5	2;0.1	1.178
6	2;0.7	1.287
7	2;0.23	1.874
8	2;1.23	1.904
9	2;2.3	1.883
10	2;2.17	2.009
11	2;3.7	2.184
12	2;4.14	2.604
13	2;5.6	2.476
14	2;5.26	2.631
15	2;9.6	2.633
16	2;10	2.465
17	2;10.18	2.735
18	2;11.12	2.873
19	3;0.3	2.690
20	3;0.17	3.306

Table I. Age and MLUw across the 20 Italian files

file	age	MLUw
1	1;10.1	1.156
2	1;10.20	1.376
3	1;11.4	1.284
4	1;11.18	1.096
5	2;0.1	1.393
6	2;0.23	1.204
7	2;2.12	1.861
8	2;2.24	1.989
9	2;4.7	2.215
10	2;4.29	1.921
11	2;7.8	2.511
12	2;9.6	2.476
13	2;10.1	2.655
14	2;10.15	2.796
15	2;10.23	2.379
16	2;10.30	2.649
17	3;0.3	2.381
18	3;0.16	2.934
19	3;1.25	2.588

Table II. Age and MLU across the 19 English files

Table III and Table IV list the recordings by file number and age, and for each recording there is a list of all participants and of the main activities.

File	Age	Participants	Activities
1	1;10.8	Elena	looking at picture books
2	1;10.27	Ludovica	looking at picture books
3	1;11.17	Ludovica	looking at picture books
4	1;11.25	Ludovica	looking at picture books, drawing, naming colours
5	2;0.1	Ludovica	looking at picture books
6	2;0.7	Ludovica	looking at picture books, drawing
7	2;0.23	Ludovica	looking at picture books, drawing, playing with coloured blocks, making a jigsaw puzzle
8	2;1.23	Ludovica, Alessandra	looking at picture books, drawing
9	2;2.3	Ludovica	looking at picture books, playing with coloured blocks
10	2;2.17	Ludovica	looking at picture books
11	2;3.7	Ludovica	looking at picture books, drawing, playing with toy animals
12	2;4.14	Ludovica	looking at picture books
13	2;5.6	Ludovica	playing with toy animals, naming parts of the body
14	2;5.26	Ludovica	preparing a snack, looking at picture books
15	2;9.6	Ludovica	playing with Lego, talking about favourite foods, playing with a teddy bear, C. trying to explain a board game
16	2;10	Ludovica	making a jigsaw puzzle, talking about nursery
17	2;10.18	Ludovica	looking at picture books, talking about nursery, using a toy telephone, C. telling L. Bambi's story
18	2;11.12	Ludovica	looking at picture books, drawing, C. telling L. Snowhite's story
19	3;0.3	Ludovica	making a puzzle, making a Lego farm
20	3;0.17	Ludovica	playing with Lego animals, C. telling made-up stories about various animals

Table III. Participants and activities across the 20 Italian files

File	Age	Participants	Activities
1	1;10.1	Karen	looking at picture books, making animal noises
2	1;10.20	Karen	naming objects and parts of the body, playing with a teddy
3	1;11.4	Karen	looking at picture books, taking out a jigsaw puzzle
4	1;11.18	Karen	playing with coloured blocks
5	2;0.1	Karen	looking at picture books
6	2;0.23	Karen	looking at picture books, making animal noises
7	2;2.12	Karen	playing with coloured blocks
8	2;2.24	Karen, Ludovica	looking at picture books, drawing, naming colours
9	2;4.7	Bob, Jim, Struan naming	toy animals, making a tower with coloured blocks, jumping and running around
10	2;4.29	Claudia, Josh, *Marco,*Ludovica	book reading, playing with coloured blocks
11	2;7.8	Sarah, Rachel, Amy *Ludovica, *Carmela	book reading, looking at picture books, running around
12	2;9.6	Eric, Ludovica	looking at picture books, playing with Lego
13	2;10.1	Eric, *Bob, *Ludovica	playing picture domino
14	2;10.15	Eric, *Ludovica	playing picture domino, looking at picture books
15	2;10.23	Eric, *Ludovica	playing with Lego, flying a toy plane
16	2;10.30	Eric, *Ludovica	playing with Lego, playing picture domino
17	3;0.3	Eric, *Ludovica	making jigsaw puzzles
18	3;0.16	Eric, *Ludovica	playing with Lego and toy animals
19	3;1.25	Eric, *Ludovica	talking to L. about friends from nursery, talking to E. about nursery, playing with Lego, making animal noises

Table IV. Participants and activities across the 19 English files

As well as the file number and age of C., a list of participants other than C. is included for every session together with a brief description of the activities that took place during the recording. The asterisk appearing before some of the participants' names indicates that although they were present at some point during the recording,

their participation was very brief, often limited to a couple of utterances per recording. As is clear from Table IV, I never fully participated in the English sessions, except occasionally at the beginning or at the end of the recording to check on the video equipment. The rationale behind this decision was twofold: firstly, I intended to keep the two languages as distinct as possible, thus reflecting the kind of distinction operating in C.'s upbringing (Italian at home and English outside the home). Secondly, my presence was inevitably a prompt for C. to start speaking Italian to me while ignoring the English speaker. This was unfortunate; it meant a decreased amount of English data which I already had objective difficulties in collecting. For the above two reasons I tried to avoid interfering with the English sessions as much as possible, especially when C. was still very young and not very talkative. By the time he was 2;9, after his return from the family summer holidays in the United States, I started spending more time with him at the end of the sessions with E., a friend who assisted me for the last six-months of the English data collection. By that time C. had become generally more communicative and talkative and we managed to have three-way conversations where E. and I would speak English to each other, he would speak English to C., I would speak Italian and C. would switch between English to E. and Italian to me thus showing a sophisticated knowledge of language choice according to the different participants to the conversation[10].

As shown in Table III above, I was C.'s main Italian conversational partner during the recording sessions, although there are two sessions where E. and A., two of C.'s Italian childminders at the time, assisted me. Both of them were familiar with C. and during that period used to see him once a week when they looked after him and his brother M..

For the English data, K., a Canadian friend of C.'s family, helped me for the first six months. She agreed to visit the house once every two weeks when she would come and play with C. for approximately 30 to 45 minutes while I recorded them. In addition to these planned visits for recording purposes, K. was a regular visitor to C.'s house and she was familiar with the child. Therefore there is reason to believe that these arranged visits for data collection purposes were not perceived by C. as an unusual event, in fact he always seemed quite happy to see K. and enjoyed playing with her. Besides the recording sessions with K., there are three sessions where C. interacts with three children of similar age from his nursery and their parents (J. and his mother;

[10] See Fantini (1978), McClure (1985) and Lanza (1992; 1997), among others, for the view that the earliest examples of systematic code-switching are a function of the category *participant.*

S. and his father; R., her mother and her sister A.). Having other children and their parents in the sessions clearly changed the dynamics of the interaction. Firstly, it was no longer a dyadic situation with one child and one adult talking exclusively to each another, but a polyadic situation where the possible exchanges included the children talking to each other, one or the other child talking to the adult, the adult talking to one or the other of the children or to both of them simultaneously, and, although a rarer occurrence, one or the other of the children addressing the other child and the adult together. However, although potentially the number of possible exchanges is greatly increased in a polyadic situation as opposed to a dyadic one, the children here did not so much focus on verbal exchanges either between them or with the adult, but tended to run around the room, jump and engage in all sorts of play activities that did not require much verbal interaction. Therefore, although rich and interesting from a more social and behavioural point of view, these recordings did not yield a very substantial amount of linguistic data.

The last part of the data collection took place in the five months after C.'s second summer in the United States. K. was no longer available, and recording with other young children had proved partially unsuccessful. I decided to return to the original format of the dyadic interaction with an adult. A friend, E., was persuaded to help me and agreed to meet C.. C. took to E. immediately, and we agreed on weekly visits when E. would come to the house and play with C. The initial lack of familiarity between C. and E. did not create any major difficulties, and with time E.'s visits became a usual and welcome habit for C.

3.8. The recording sessions: the setting and the activities

3.8.1 The setting

With the exception of four recordings which took place either in C.'s bedroom or in M.'s, the rest of the sessions took place in the large sitting room in C.'s parents' house. The room was very big and in the middle of it there was a big rug where C. could sit and play with his toys and look at his books. The room had a big bay window and plenty of natural light was always available. The videocamera would normally be balanced on a chair or on a tripod so that a good portion of the playing area would be in view to include C., his partner(s) and the books/toys he used.

The recordings usually took place in the early afternoon after C. had returned from nursery and had had the opportunity to play with me or the childminder in charge for half an hour to an hour at home. During this time of day he was usually quite active and willing to play, while, especially when he was younger, he tended to become tired and quieter towards the middle-end of the afternoon. Also this was the time when C. was home alone with his childminder and it was less likely to be interrupted by M. or visitors coming to the house. Although M. participated to a few sessions, I did not normally include him since, because of his much higher verbal skills, he had a tendency to monopolise the adult's attention making it more difficult to include C. in the conversation.

3.8.2 The activities

Tables III and IV list a number of activities that took place during the various recordings. It must be noted that while the sessions were unstructured and there was no pre-arranged set of toys, books or activities that would be used in any particular session, the choice of material was of course not random and tended to be age-appropriate. In other words, the kind of toys and activities in the recorded sessions reflected C.'s interests and abilities at that particular moment. Of course, some of the props, such as books, remained the same throughout the data collection period, although the use that was made of them changed over time. In the case of books for example, they were initially a visual support for naming routines where the adult opened with a question such as *What is this?* or *Where is X?* and C. was expected to answer either by providing an appropriate label, or by pointing and/or using a linguistic deictic device. Then the adult either praised the child for the correct answer by confirming it in a number of ways (repetition, repetition with expansion, recast, praise, e.g. *Yes, that's right* or *Well done,* etc.), or alternatively signalled that his answer was not appropriate and would either recast or repeat the question until the child answered correctly, or, depending on the child's responsiveness, moved on to something else. Below are two of many examples of these naming routines that abound throughout the early English and Italian files:

(1) File carloeng1.mor

*KAR: what's that?
*CAR: raffe.
*KAR: what's that?
*CAR: elepha.
*KAR: mmhm mmhm.

(2) File carloita1.cha

*CAR: ina.
%eng: small.
*ELE: ina è una tartarughina.
%eng: small it's a small turtle.
*ELE: e vicino alla ina cosa c' è?
%eng: and what's next to the small one?
*CAR: ina.
%eng: small.
*ELE: cos' è che ha qui la tartarughina?
%eng: what's the turtle got here?
*ELE: cos' è che ha sopra la testa?
%eng: what's it got on its head?
*CAR: testa.
%eng: head.
*ELE: come si chiama questo lo sai?
%eng: do you know what this is called?
*CAR: ina.
%eng: small.
*ELE: no si chiama cappello.
%eng: no that's called hat.

Over time the books' function changed and was no longer a mere support for labelling routines, but for more complex descriptive tasks or primitive narratives:

(3) File carloita10.cha

*LUD: dove va questo coniglietto?
%eng: where does this bunny go?
*CAR: hmm dove va?
%eng: hmm where does it go?
*CAR: sul buco.
%eng: on the hole.
*LUD: va nel buco e questi coniglietti cosa fanno?
%eng: it goes into the hole and what do these bunnies do?
*CAR: correne.
%eng: they run.
*LUD: corrono.
%eng: they run.
*CAR: corrono.
%eng: they run.

(4) File carloeng8.mor

*CAR: da(t) Aladdin .
*KAR: that's Aladdin .
*CAR: what doing ?
*CAR: what doing ?

By the end of the recordings, the books had become almost a pretext to elicit more articulated story telling:

(5) File carloeng8.mor

*ERI: what's happening here ?
*CAR: what's he doing ?
*ERI: what are they doing .
*CAR: ahem <they're bringing> [/] ahem they're bringing this .

(6) File carloita17.cha

*LUD: cosa succede nella storia di Bambi durante il temporale?
%eng: what happens in Bambi's story during the storm?
*CAR: è morta.
%eng: she is dead.
*LUD: chi è morta?
%eng: who is dead?
*CAR: 0 [=! mumbles].
*LUD: hmm?
*CAR: la mamma.
%eng: the mummy.
*LUD: ahhah e chi l' ha uccisa?
%eng: ahhah and who killed her?
*CAR: ahem il temporale.
%eng: ahem the storm.
*LUD: il temporale?
%eng: the storm?
*CAR: sì.
%eng: yes.
*LUD: no # non l'ha uccisa il temporale.
%eng: no # the storm did not kill her.
*LUD: chi l' ha uccisa?
%eng: who killed her?
*CAR: ahem ahem la neve.
%eng: ahem ahem the snow.
*LUD: no!
%eng: no!

Other activites include playing with coloured wooden blocks, drawing, making Lego constructions, talking on a toy telephone, assembling jigsaw puzzles, playing picture domino. From a very young age C. displayed a keen interest in games that involved some sort of assembly, whether it be building a tower with coloured wooden blocks, pieces of Lego, or making a jigsaw puzzle. Together with looking at picture books, playing with coloured wooden blocks was from the beginning one of C.'s favourite activities which was usually very suitable to try and elicit colour names and size terms when comparing blocks of different colours and sizes.

(7) File carloeng7.mor

*KAR: that is blue .
*CAR: that yellow ?
*KAR: that's yellow .
*CAR: that blue !
*KAR: that's blue .

Drawing was another activity that proved useful in getting C. to talk about what he was doing. He would start by either drawing something himself, or by asking the adult to draw something for him, and then together they would interpret the drawing. It was normally some animal figure and the game consisted of trying to identify the various body parts and colour them in:

(8) File carloita18.cha

*CAR: c' hanno le corne?
%eng: do they have horns?
*CAR: c' hanno le corne?
%eng: do they have horns?
*LUD: i cammelli hanno le corna?
%eng: do camels have horns?
*CAR: sì.
%eng: yes.
*LUD: dici?
%eng: do you think so?
*CAR: guarda il suo naso.
%eng: look at his nose
*LUD: quello è il naso del cammello?
%eng: is that the camel's nose?

(9) File carloeng8.mor

*CAR: a dog ?
*KAR: it's not a mouse .
*KAR: it's just a dog with a big nose and big ears .
*CAR: a cat .

Building constructions with pieces of Lego became one of C.'s favourites around his third birthday, especially during E.'s visits. Although at times C. seemed to focus exclusively on the building task and was unwilling to engange in verbal exchanges with E., making trains for the toy animals, building tunnels and houses was a source of rich interaction with his adult partner. With E. in particular, playing with Lego would leave him the opportunity to say nothing if he so wished, without having to interrupt the game. In the case of more demanding activities, such as looking at

picture books or telling stories, in order not to interrupt his contact, C. would have to talk and respond appropriately. When unwilling to comply with the demands imposed by his interlocutor, C.'s strategy was either to remain silent or to reply "I don't know":

(10) File carloen14.mor

*ERI: what happened, Carlo ?
*CAR: I don't know .
*ERI: <I don't know> ["] ?
*ERI: I think you do .
*ERI: what can we see in the sky ?
*CAR: 0.
*ERI: what are they ?
*CAR: I don't know .

The practical task involved in making Lego constructions would instead give him the excuse to go on with the game without necessarily having to talk. He was still being cooperative nevertheless, and was contributing to the joint task at hand, even if he was not actually saying anything. Moreover, the very nature of the game, involving the creation of several possible made-up scenarios (animals and people boarding a train or a plane; animals feeding; people and animals building a house/a farm/a station/a tunnel, etc.) gave C. and his partner the opportunity to talk about a variety of situations in a more open-ended way with respect to book reading, where the situations were by and large already set, and it was more a question of simply describing them, rather than making them up.

If at one end of a continuum we can place playing with Lego as one of the most productive situations in which C. was at its most spontaneous and creative, at the opposite end of the continuum we find activities such as talking on a toy telephone and playing picture domino which created the largest amount of formulaic language and routines. C. seemed to perceive these games as very predictable and highly structured, and as a consequence would follow a routinised pattern where he repeated "his lines" over and over again, almost as if rehearsing a script that was always the same.

(11) File carloeng16.mor

*CAR: and one for <Carlo> [>] .
*ERI: <Carlo> [<] .
*CAR: and one for Eric .
*ERI: yeah .
*CAR: and one for Carlo .
*ERI: ha ha ha !
*CAR: ha ha ha .

*ERI: 0 [= laughs] .
*CAR: and one for Carlo .
*ERI: one for Carlo .
*CAR: and one for Eric .

Although the tendency to follow a predictable pattern of routinised behaviour, whether it be systematically avoiding compliance with the interlocutor's requests or repeating a formula over and over again, can be identified in both language contexts, there is evidence that it is significantly more pronounced in English than in Italian. A more fine-grained analysis of the data reveals a difference in C.'s conversational competence in the two languages. In English the number of C.'s verbal turns is lower, he often resorts to non-verbal behaviour to acknowledge a previous utterance from another interlocutor, or simply ignores it altogether. The number of repeated utterances, either self-repetitions or repetitions of a previous utterance by another speaker, is considerably higher in English. The number of routines and formulas is also greater in English as compared to Italian.

The picture that emerges from this brief overview of the activities during the various recorded free-play session is one where C. displays a variety of communicative strategies when interacting with his various partners. Not only are these strategies predictably age-dependent, but, to a certain extent, they are also activity-dependent and language-dependent. There is clearly a gradual progression in the extent to which C. engages in conversation with other speakers: over time the number of C.'s turns increases; he learns to comply with the speaker's expectations by providing appropriate answers, engaging in scaffolding routines, taking over the initiator's role more and more often, asking questions rather than simply answering them, etc. This is of course not surprising, it is only to be expected that as the child grows older his communicative competence gradually becomes more sophisticated. What is interesting here is that his competence is expressed in different ways according to the situation he is in, i.e. what type of activity he is engaged in and what language is being used. Some activities, such as playing picture domino, tend to elicit a form of routinized and stereotyped behaviour, while others, playing with Lego, seem more open-ended and foster a greater deal of more spontaneous speech.

3.9. Transcription and coding

The videorecordings used in this study were entirely transcribed by myself using the standardized CHAT conventions as described in MacWhinney (1995)[11]. CHAT (Codes for the Human Analysis of Transcripts) is the transcription system used on the CHILDES (Child Language Data Exchange System) Project, and is a standardized format for computerized transcripts of face-to-face conversational interactions. The creation of the CHILDES project has proved of vital importance for child language researchers all over the world by making it possible to share data in a readily accessible, standardised and computerised format. Thanks to the invention of the CHAT transcription system, we are now a long way from the situation described by Elinor Ochs in 1979: "[o]n a very practical level, we have not developed a set of conventions for representing the verbal and nonverbal actions of young children" (Ochs, 1979: 44). Now we do have this set of conventions for transcribing child data, and it was with the intention of producing transcripts that could be of use to the research community at large that I chose CHAT for my data.

The data was transcribed orthographically except for some child forms for which a broad phonetic transcription was provided together with the corresponding adult target. In addition to the the speakers' utterances transcribed on the main line, a %mor dependent tier was also created for the morphological tagging of the English data. An %err tier for the coding of errors, and a %gpx tier for the coding of gestural and proxemic material were also used whenever necessary. A detailed description of the transcription and the coding systems adopted is provided in Appendix 2.

3.10. Summary

This chapter has introduced a definition of bilingual first language acquisition (BFLA) as the regular exposure to two languages from the first week of life (De Houwer, 1990). Some possible bilingual input patterns have been described in connection with the selection procedure for the informant of this case study, and a number of factors have been identified as concomitantly responsible for successful childhood bilingualism: balanced access to both languages with particular emphasis on the minority language, adherence to

[11] It is unfortunate that the transcripts could not be double-checked by a second transcriber, but at the time I could not find anyone who was proficient in both languages, had some experience of transcribing child language and was also familiar with the CHAT system.

the one-person one-language strategy, and a positive parental attitude to multilingualism. In the case of Carlo, the English-Italian bilingual child chosen for this study, the family circumstances were ideally suited to the child's acquisition of both languages: both his parents speak English and Italian and have a very positive attitude towards bilingualism; C.'s older brother M. is a very successful bilingual himself, exposure to Italian is maximised by the employment of Italian-speaking babysitters, visits to the maternal grandparents in Italy, and C.'s father speaking Italian to the child.

An in-depth description of C.'s linguistic environment has been provided here for both languages, and observations at the two nursery schools the child attended has provided valuable information on differences on the nature of the input at home in Italian, and outside the home in English. Information was also provided on the recording sessions and on the transcription and coding system used.

The following chapter introduces the theoretical framework adopted throughout this work, and provides a critical review of current approaches to language acquisition with particular reference to the acquisition of morphosyntax.

Chapter 4

Verbs, Features and Functional Categories in Early Child Grammar

4.1. Verbs and formal grammatical features

The main focus of this thesis is on the grammatical development of an English-Italian bilingual child with particular reference to the emergence of Functional Categories (FCs) in the child's two languages. The first part of this chapter gives a critical overview of a number of approaches that have appeared in the language acquisition literature on the emergence of the Verb category and its associated formal grammatical features. The second half introduces a number of approaches to language acquisition and the emergence of verbal syntax.

Verbs have a privileged role in language acquisition because of their centrality to the expression of grammatical relations. Verb knowledge entails a complex set of semantic, morphological and syntactic information: categorial status, argument structure, meaning, morphophonological properties of the corresponding Agreement , Tense and Mood and Aspect paradigm, etc. For instance, knowing the lexical entry for the Italian word *parlare, "to talk"*, means that one knows that it is a verb, and specifically that it is an unergative verb optionally taking a dative PP complement (*parlare (a qualcuno), "to talk (to someone)"*), that it belongs to the *-are* conjugation, that it is a regular verb and that, as such, will take the appropriate Agreement, Tense and Mood affixes of the *-are* conjugation.

The extent to which children master the argument structure of a verb and its grammatical features is typically taken as an indication of their mental representation of syntactic/functional categories. One of the central issues in language acquisition research is how early, and how broadly, this abstract knowledge about functional features is instantiated in children's grammars. The approaches to the emergence of syntactic representations overviewed here vary along a continuum: at one end we find structure-building approaches that deny the existence of FCs in Early Child Grammar (ECG) altogether (Radford, 1990 etc), at the other end we find proponents of continuity

approaches that argue for the availability of FCs as a prerequisite of language acquisition, thus minimising discontinuity between adult and child grammar. Somewhere in the middle ground we find lexicalist approaches proposing that, altough FCs are in principle available to children and thus not subject to any maturational constraints, however their realisation is driven by the learning of lexical items and their associated grammatical features. Lexicalist approaches include a variety of positions, ranging from those set in a generative grammar framework (Clahsen et al.), to more constructivist approaches such as those proposed by Tomasello (1992), and Pine, Lieven and colleagues (lexical distributed learning).

In section 4.2 I will give an analysis of the status of functional features and FCs in the latest version of Chomsky's syntactic theory, the Minimalist Program (MP); section 4.3 will look at how basic tenets of the Chomskyan syntactic model have been applied to the central issue of the mental representation of functional verbal features. Section 4.4 is devoted to more lexical-constructivist approaches as proposed by Tomasello (1992), and more recently by Lieven, Pine & Baldwin (1997), Pine & Lieven (1997), Pine & Rubino (1998), and Pine, Lieven & Rowland (1998). Section 4.5 compares some of the positions overviewed here and proposes to reconcile some aspects of Chomsky's theory with a more constructivist apporach that will take into account the verb-specific organization of children's grammars and lexically distributed learning.

4.2. Formal Features and Functional Categories in the Minimalist Program

4.2.1. Interpretable and -Interpretable features

The functional information carried by the grammatical features encoded in every item's lexical entry plays a vital role in the justification of syntactic models and in approaches to language acquisition based on such theoretical models. In generative syntactic theory, and more specifically in Chomskyan models of syntactic theory, the functional information carried by verbs in the form of inflectional affixes has abstract correlates in the form of Functional Categories. Among the features that are included in lexical entries, Chomsky (1995: 230) distinguishes "between *formal* features that are accessible in the course of the computation and others that are not", those features that are not accessible to the computation are purely semantic features, e.g. the feature

[artifact] for the lexical entry for *book*.[1] Formal features can either be intrinsic to the lexical item or added optionally when the item is selected in the Numeration, where a Numeration is intended as a set of pairs (LI, *i*), in which LI is a lexical item, and *i* is the number of times that LI is selected. Take, for example, the sentence in (1):

(1) I love dogs.

Intrinsic features of the lexical items include the categorial feature, whether an item is a Determiner (*I*), a verb (*love*), or a noun (*dogs*); [1 person] in FF(*I*), [3 person] in FF(*dogs*), assign Accusative Case in FF(*love*) and assign Nominative Case in FF(T). Optional features that are selected in the Numeration process are [plural] for FF(*dogs*), and the phi features [3 person plural] for FF(*love*).

A further crucial distinction exists within formal features between -Interpretable and Interpretable features. Only Interpretable features, but not -Interpretable features receive an interpretation at the Conceptual-Interpretive interface. In Chomsky (1995) Interpretable features include categorial features, phi features of nominals, and Tense; -Interpretable features are typically Case, and phi features of verbs.[2]

The working hypothesis of the MP is that instead of S-structure and D-structure, there are only two interface levels in the grammar: PF (Phonetic Form), or the Articulatory-Perceptual interface (A-P), and LF (Logical Form), or the Computational-Interpretive (C-I) interface. Through an operation called Select, lexical items are taken from the lexicon to enter the Numeration from which constituents are combined together in a pairwise fashion through a process of Merge. At a certain stage in the derivation, the level of Spell Out, the phonetic and semantic features are processed separately, the former being interpreted at the A-P interface by PF operations, and the latter being interpreted at the C-I interface by LF operations. A UG constraint known as the Principle of Full Interpretation requires that a representation for a given expression must contain all and only those elements which contribute directly to its

[1]Chomsky however admits that even purely formal features, such as transitivity and accusative Case for instance, reflect semantic properties (Chomsky, 1995: 381, fn 14).

[2]Phi features of Greek nominals have been treated as -Interpretable by Tsimpli & Stavrakaki (1999). It is not the phi features on nouns as such that are considered -Interpretable at LF, however, but rather the resumptive features on the heads of the nominal projection which result from a head-dependency head-formation between them. Thus phi features on the N head are represented on all heads participating in the Chain, including adjectives and modifiers, in the form of resumptive features. Such resumptive features are erased at LF as -Interpretable, while the heads in the chain either survive or not depending on their Interpretability. Thus N, because of its categorial features, and Def because of its Interpretable [definiteness] feature are not erased because Interpretable, while D which is only specified for Case (-Interpretable) and resumptive features (-Interpretable) is erased.

interpretation at the relevant level: thus the only features that the PF representation of an expression can contain are those relevant to determining its Phonetic Form, and in a similar fashion the only features that are allowed in a LF representation are those that are relevant to determining its logical form. When the phonetic and semantic features of an expression are represented at the appropriate interface levels, a derivation is said to converge. If, on the other hand, features are illegally represented at the inappropriate level, e.g. phonetic features appear in a LF representation, the derivation for that expression is said to crash.

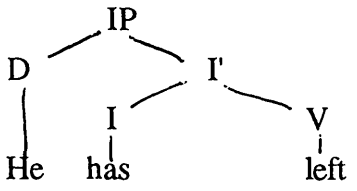
If a derivation is to converge at LF, -Interpretable features must then be somehow eliminated in the course of deriving an LF representation. The process whereby -Interpretable features are erased is checking. Once checked, -Interpretable features are inaccessible to the Computational System and are invisible at LF. By constrast, when Interpretable features are checked they are not erased, they are still accessible to the Computational System and are visible at LF. Thus for example, once Case [-Interpretable] is checked, it is erased and hence invisible to LF derivations and inaccessible for further computations. Conversely when Interpretable features such as phi features of nominals are checked, they are nonetheless accessible for further computation and still visible to LF processes. The implication of checking of -Interpretable features as erasing is twofold: firstly, if all -Interpretable features of a category a have already been checked by an appropriate target, the movement operation known as Move α is blocked. Move α only applies as a last resort operation when checking cannot take place in a given configuration. Secondly, because only -Interpretable features are erased after checking, a category a can enter into multiple checking relations for Interpretable features (e.g. categorial features and phi featues of a DP), but multiple checking of -Interpretable features is blocked by erasure (e.g. the Case feature of a DP can only be checked once, after which it is no longer available for successive computations).

4.2.2. Checking mechanisms

In concrete terms, feature checking takes place in a relevant spec-head configuration which can be either the result of a Merge operation or of a Move operation. Merge is an operation by which two categories are combined to form another category, the resulting phrase being endocentric, i.e. it is the projection of a head. Heads select specifiers and complements, and the range of lexical items that can appear as specifiers and complements of a given head to form new constituents is

determined by the specifier and complement features carried by the relevant head. The assumption here is that lexical items carry three different sets of grammatical features: head features (which determine their intrinsic grammatical properties), specifier features (which determine the kind of specifiers they allow), and complement features (which determine the kind of complements they select). A head checks features of its specifier and its complement, all specifier and complement features are -Interpretable, as are purely formal head features (features with no semantic content, hence invisible at LF). Consider (1) and how the process of feature checking works:

(1) He has left.



[3MSNom]	[Pres]	[+n]	head features
	[3SNom]		specifer features
	[+n]		complement features

For each of the three lexical items in (1) we have categorial features, i.e. we know whether they are a Determiner (*He*), an auxiliary (*has*), or a verb (*left*). Moreover for each lexical item there is a specification of their head features, 3rd Person, Singular Number, Masculine Gender (phi features), and Nominative Case for *He;* Present Tense for *has*, and n-participle form for *left.* Because neither *He,* nor *left* have a specifier or a complement in (1), specifier and complement features are not listed.

In the checking process, the specifier features of the head are checked against the head features of its specifier, and similarly the complement features of the head are checked against the head features of its complement. If there is compatibility between checker and checked with respect to a given feature, the relevant specifier or complement feature is erased, since specifer and complement features are always -Interpretable. The corresponding head feature is erased only if -Interpretable, if it is Interpretable after checking it is not erased so that it can still be accessed by the Computational System and can remain visible for LF processes. In (1) the specifier features of the head *has*, [3SNom] are checked against the head features of the specifier *He*, [3SMNom], because there is compatibility between the two sets of features, the specifier phi features of the head are erased, being -Interpretable, while

they are checked but not erased for the specifier *He*, since phi-features of nominals are Interpretable, and after checking they must remain available to the Computational System for LF processes. As for Nominative Case, it is erased both ways, being always a -Interpretable feature.

As for the checking between the complement feature of the head, [+n], and the head feature of the complement, [+n], there is compatibility and the feature is checked and erased both for the head and the complement, on the assumption that inflectional properties of non-finite verbs play no role in semantic interpretation and are therefore -Interpretable. The only formal grammatical features that survive the checking process and are not erased are interpretable head features: [3MS] for the specifier *He,* and [Pres] for the auxiliary *has*. These must remain accessible to the Computational System for LF processes, so that the derivation can converge both at the PF interface, through erasure of -Interpretable features, and at LF through checking of the remaining Interpretable features.

4.2.3. Move α and Move F

As briefly sketched above, feature checking must take place in an appropriate checking domain. When an item is not in such a checking configuration, an operation called Move ensures that the given item enters in the required configuration with an appropriate target. The operation Move creates a chain CH = (α, t(α)) which must satisfy three fundamental conditions: the uniformity condition, whereby a chain must be uniform with regard to phrase structure status (an element must maintain its phrase structure property of being minimal, maximal or neither); α must c-command its trace, (any Move operation is a raising operation); finally a chain CH is formed only if Move satisfies a Last Resort condition.

The assumption is that when Move applies, a category α is raised to target a category K, both α and K being constructed from one or more lexical items. However, since Move is justified by feature checking, the minimal assumption should be that only the relevant feature F, and not the whole category α, is actually being raised to target K: the operation Move should apply to a given lexical feature F, and not to a category α. Take, for instance, the case of a Subject DP raising to spec-AGRS to check its Case feature, why could it not be the case that only the relevant feature Case actually raises to AGRS to be checked, as opposed to the whole DP? Chomsky (1995: 262) proposes (3) below:

(3) F carries along just enough material for convergence.

Although the operation Move really only aims at raising the feature F that needs to be checked, properties of the A-P system require that extra phonological material be pied-piped to satisfy PF requirements. Failure to comply with these PF constraints would cause the derivation to crash at the PF interface, and the Principle of Full Interpretation would be violated.[3] It follows that PF conditions, and the pied-piping of extra phonetic material, will only be relevant in the case of overt raising. What happens then in covert raising? Does the feature F of the lexical item LI actually raise alone, or does it also take other formal features along with it? Chomsky suggests that there are stong empirical reasons for supposing that Move F automatically carries along other formal features of the relevant lexical item (Chomsky, 1995: 265):

(4) Move F "carries along" FF[F]

On this assumption it is therefore the case that Move always creates an additional chain $CH_{FF} = (FF[F], t_{FF[F]})$ alongside the chain formed by movement of the feature F, $CH_F = (F, t_F)$, and in the case of overt raising a third chain is formed: $CH_{CAT} = (\alpha, t\alpha)$. What is of interest to the Computational System is really only CH_F, but the Computational System can also see the other two chains, whose *raison d'être* is simply to satisfy a PF requirement in the case of CH_{CAT}, and the inclusion of "free riders" in the case of CH_{Ff}, additional formal features which are not directly targeted by movement for checking purposes.

Move F has been defined as movement of a feature F, and possibly of additional formal features FF[F], or even a whole category α, to a target K. This operation is justified by the morphological requirement to check, and erase in the case of -Interpretable features, the relevant feature F. An additional condition on Move is that movement of F to K is allowed only if some feature of K is available to enter into a checking relation with F. In other words, the raised unchecked feature F must enter into a checking relation, which is only possible if the target K has an as-yet-unchecked feature.

In more concrete terms, take again the example of the checking of the -Interpretable Case feature on a Subject DP. Overt raising of the whole DP will move

[3] Chomsky observes how it is the imperfection of the sensorimotor apparatus that imposes the inclusion of extra phonetic material. The Computational System would in fact require a more minimal operation simply involving the raising of features and not of whole lexical items.

the Case feature, along with Interpretable phi features as free riders, to spec-AGRS. This operation is legitimate on the assumption that the target, in this case AGRS, has not yet checked the Nominative Case feature; if this is so the raised DP will check the relevant Case feature in an appropriate spec-head configuration. Being -Interpretable, the Case feature will be erased after checking and will no longer be accessible either to the Computational System or to LF processes, thus ensuring PF convergence. The free riders nominal phi features will also be checked, but being Interpretable they will not be erased and will be visible to LF processes. At the same time, phi feature checking will erase the corresponding phi features in AGRS, since phi features of verbs are -Interpretable and need to be erased after checking.

4.2.4. The status of functional categories

Because of their role in feature checking, functional categories (FCs) are of crucial importance in the Move/Attract process. The inventory of FCs considered in Chomsky (1995) includes T, C, D, and Agr. The first three carry Interpretable features and are thus visible at the LF interface. T is [± finite], with a further subdivision [± Past], thus providing information about temporal reference; D is identified as [± referential], while C is an indicator of illocutionary force, specifying whether a given clause is declarative, interrogative, and so on.

Agr, on the other hand merely checks -Interpretable Case and phi features of predicates which have no relevance at the LF interface. Moreover, Chomsky argues that Agr does not independently carry Case and phi features, Case features are provided by the V that adjoins to it; so are phi features which are optionally added to the verbal predicate when it is selected from the lexicon for the Numeration. If Agr's only reason to be is to check -Interpretable features that it does not even carry independently, Chomsky (1995: 349 ff.) argues that the postulation of such a functional category is justified only by theory-internal reasons and is not necessary on more principled grounds. He thus proposes to dispense with Agr projections altogether (both AgrS and AgrO), and posits that Case and phi features of predicates should be checked by T. In addition T would also have a strong nominal D feature requiring the movement of a Subject DP in its spec position, and depending on whether the strong nominal feature of T is such that, although it is -Interpretable, it is not erased after checking, T can have multiple specifiers as in the case of Multiple Subject Constructions (MSCs) in Icelandic. Whether the strong nominal feature of T erases after checking or not would be subject to parametrization. Therefore in a language like English where MSCs are not

allowed the -Interpretable nominal feature erases after checking, and only one DP can end up in spec-TP. In languages like Icelandic where MSCs are allowed the parameter of the strong -Interpretable feature would be set so that checking does not erase the feature after checking and more than one specifier is allowed in TP. As Cann (p.c.) points out, however, the notion that erasure of -Interpretable features is subject to parametrization critically detracts from the very definition of -Interpretable features as features that are erased after checking.

In Chomsky's revised approach to FCs, only those with intrinsic Interpretable features survive. T and V also carry -Interpretable features that need to be checked: for T [(assign) Case] (nominative or null); for V, its phi features and [(assign) accusative Case]. In this respect phi features of verbs are like phi features of nouns in that they are optionally selected when the lexical item is drawn from the lexicon for the Numeration.

Van Gelderen (1993: 145) proposes a Subject Agreement Parameter, whereby the realisation of an AGR category in a given language would be justified by relevant morphosyntactic evidence:

(7) Subject Agreement Parameter: ±AGRs/AGRsP

For English she argues that there is no convincing evidence for a second functional projection beside T, between C and NEG. The first piece of evidence comes from the well-known fact that no element can intervene between C and a modal or infinitival *to* in T:

(8)a. *that he may to work.
b. *that he may can work.

Van Gelderen claims that (8)a and (8)b are evidence that there is no room for another functional node higher than the T node hosting *to* and the modal. I would argue that what (8) simply shows is that *may* competes for the same structural position with another modal in (8)b, and with inifinitival *to* in (8)a.

(9) *The kids may all not go there.

Another alleged problem is posed by sentences like (9) above. The subject DP *the kids* is assumed to be in spec-AGRSP, the modal *may* has moved to AGRS, the

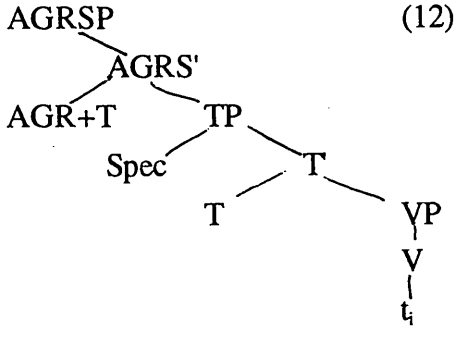
floated quantifier *all* is therefore in spec-TP, and *not* is in NEG. By the ungrammaticality of (9), van Gelderen concludes that only one specifier position can be projected: spec-TP, where the subject DP is allowed to raise. More recently, Jonas (1994), Bobalijk & Jonas (1996), and Bobalijk (1997) have in fact argued that spec-TP as a subject landing site is subject to parametrization, and that English instantiates the negative value of the parameter, it is a [- spec-TP] language:

(10) The Spec-TP Parameter
Some languages license spec-TP as a potential landing site for the subject NP; other languages do not license this position.

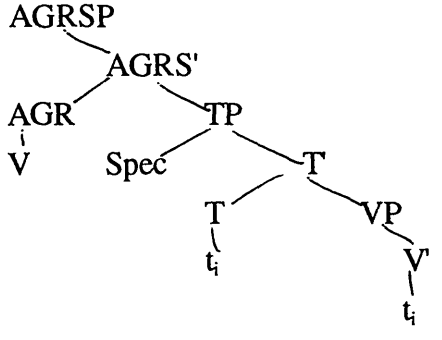
(Bobalijk, 1997: 1033)

Languages such as German, Dutch, Icelandic, Yiddish, and Afrikaans that allow object shifts of full DPs and transitive expletive constructions must be [+spec-TP] languages, while languages such as English, Danish, Norwegian and Swedish that do not, are [-spec-TP] languages.[4] Jonas argues that for feature checking with T in the overt syntax, T must be lexicalised either through the inflected verb raising to T prior to Spell Out, or through raising of T to AGRS. Languages where the verb raises overtly, lexicalising T in situ, are [+spec-TP] languages where a spec-head relationship may be established in TP. By contrast, languages where the verb does not raise overtly (with the exception of V2 phenomena), are [-spec-TP] languages where T adjoins to AGR in order to be lexicalised, and spec-head relationship is established in AGRSP. The trace of T in TP is not lexicalised, hence spec-TP is not a possible position for feature checking between a subject or expletive DP and T. The diagrams in (11) and (12) represent the [-spec-TP] option and the [+spec-TP] option respectively:

(11)



(12)



[4]For a full exposition and justification of the Spec-TP Parameter, see Bobalijk & Jonas (1996) and Bobalijk (1997).

The diagram in (11) represents the situation of a language where a single affix expresses both Agreement and Tense, or differently phrased, vocabulary items expressing Agreement or Tense are in complementary distribution. This is the case of English where Tense and Agreement morphology are in complementary distribution. The inflectional paradigm for the Present Tense is only marked for Agreement, and specifically for Number, the *-s* inflection simply distinguishes singular from plural forms. Kayne (1991) argues that

> [i]f one takes *you* to always be grammatically plural, somewhat as French vous, despite sometimes referring to a singular, then, if one takes *I* to be non-singular (there is clearly nothing that *I* is a true singular of), English -s can be considered to be a pure indication of number (+singular), rather than involving person in any way.

Van Gelderen (1993) adds that, historically, *you* is certainly a plural accusative form, while *thou* is singular nominative, *thee* singular accusative, and *yee* plural nominative. Moreover, in many languages first and second person singular are not distinct from the first and second person plural. This reduces the only Agreement feature in English to Number. Hoekstra & Hyams (1995) also argue that in English finiteness in the Present Tense is only marked by the Number feature.The Past Tense inflectional paradigm, on the contrary, is only marked for Tense, there are no Person or Number distinctions. Therefore, when an English verb is selected, it is either marked for Agreement or Tense, but never both; in this sense Agreement and Tense forms are in complementary distribution as claimed by Bobalijk (1997).

For languages where spec-TP is available as a potential landing site for the subject, (12), Agreement and Tense morphology is not in complementary distribution, but there is evidence that different affixes express different features. Observe the paradigm in (13):

(13) German *sagen* ("to say")

Present		Past	
Ich sag-(e)	wir sag-en	Ich sag-te	wir sag-te-n
du sag-st	ihr sag-t	du sag-te-st	ihr sag-te-t
er/sie sag-t	sie sag-en	er/sie sag-te	sie sag-te-n

It is clear that Tense and Agreement are not in complementary distribution, the Past paradigm has different affixes to mark Tense (*-te*) and Person/Number Agreement

(*-st, -t, -n*). In German, like in Dutch, Icelandic, and Mainland Scandinavian languages, positive evidence for distinct Agreement and Tense morphology is only available in the Past Tense paradigm, whereby Bobalijk (1997: 1044) concludes that

> [i]f the appearance of tense morphology blocks the appearance of agreement morphology, then tense and agreement vocabulary items are in complementary distribution and T and AGR must be fused.

The German paradigm in (13) also shows that in this language both Person and Number are relevant Agreement features, both in the Present and in the Past Tense. There is a two-way distinction between singular and plural number, both in the present and in the past paradigm. In the present tense singular paradigm we have a three-way person distinction, which in the plural is reduced to a two-way (no person distinction between 1 p.p. and 3 p.p.). In the past tense, there is only a contrast between 2 and non-2 person, but there still is a number distinction between singular and plural. Although Agreement will be normally employed throughout this work to include both person and number features, the distinction between these two sets of features will be specified wherever appropriate.

To sum up, FCs are the targets to which grammatical features F or whole categories, depending whether PF conditions apply, are raised to so that they can be checked as appropriate. If a -Interpretable feature F is strong in a language A, then Move will apply to the category a carrying F, a will move to target the relevant FC to check and erase the -Interpretable feature F. If, by contrast, a -Interpretable feture F is weak in a language B, Move will not need to apply and only the relevant grammatical feature F, and no extra phonetic material, will be attracted to the relevant FC.

From what has been outlined above, it is clear that the availability of FCs plays a crucial role in all the operations of the Computational system, with obvious implications for language acquisition. UG mechanisms available to the child will inform her that features need to be checked, in particular that strong -Interpretable features require the raising to an appropriate target FC where they will be erased after checking. The child is also supposed to be sensitive to the UG Principle of Economy whereby Move is a Last Resort mechanism driven by morphological properties, and that covert movement is to be preferred over overt movement. Feature checking is a two-way process, on the one hand the child must be aware that a given lexical item carries a certain number of features, secondly an appropriate FC must be available in her mental representation where the feature checking can take place either by substitution or adjunction.

Recent research in the field of language acquisition in a generative framework has focused on the crucial question of whether and to what extent FCs are available to the child in the course of acquisition. A number of proposals exist in the literature investigating the status of FCs in the initial stages of children's grammars. In section 4.3 I will present three major hypotheses on the initial state of child grammar and the access to FCs: the Structure-building Hypothesis (SBH), the Continuity Hypothesis (CH) and the Maturational Hypothesis (MH).

4.3. Three hypotheses on the initial state

4.3.1. The structure-building hypothesis

Starting from a vast empirical basis of English child language, Radford (1990, 1992, 1994, 1996) has extensively argued on a number of occasions for a Structure-building approach to language acquisition. The main theoretical assumption of Radford's approach, also shared by a number of other researchers in the field (notably Guilfoyle & Noonan (1988), Lebeaux (1987), Platzack (1990, 1992), and more recently Tsimpli (1996)) is that Early Child Grammar (ECG), and specifically Early Child English, is characterised by the absence of FCs. The inital state of language acquisition is thus a Prefunctional stage where FCs are not accessible to the child; only subsequently do they become available as the result of maturation.[5]

At the onset of acquisition, only lexical categories such as N, V, A, and P are part of the child's grammar, while functional categories such as D, Agr, T and C are not.[6] The notion of a two-tier system in acquisition, where lexical categories emerge before functional categories, finds substantiation in recent claims of linguistic theory where FCs are assumed to be the locus of all parametrization. According to the Functional Parametrization Hypothesis (Chomsky, 1989, 1995), languages of the world would differ with respect to which FCs they instantiate and to the strength of the

---

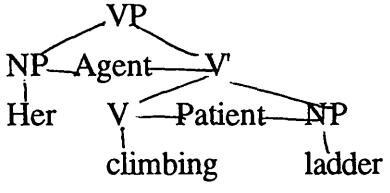
[5] Radford (1992: 25) refers to Early Child English as the period of "*early patterned speech* which typically lasts from around 20 to 24 months of age (±20%)". Because of great individual variation among children, rather than using a chronological measure of development in terms of the child age, the position adopted here is to use MLU (Mean Length of Utterance) instead. In line with Brown's (1973) criterion of identification of different developmental stages on the basis of MLU, ECG will correspond to Brown's Stage I and II (MLU between XXX)

[6] Prepositions have a dual status in that they can function both as lexical and as functional elements. Radford's approach would therefore predict that whenever prepositions play the role of a functional marker they should not appear in the earliest stages of acquisition (e.g. genitive case marking in "a piece *of* cake").

associated features. Although UG makes available an inventory of FCs, it is not necessarily the case that all of them are actually part of the grammar of each and every natural language. For example Fukui (1986) argues that Japanese differs from English, and many other languages, in not having a C system, an AGR system and a D system. Moreover, even when a FC is represented in a given language, the strength of the relevant features is subject to parametrization. Well-known syntactic differences between English and French as far as verb movement is concerned, have been ascribed to the relative strength of the AGR category in the two languages (Pollock, 1989): weak AGR features in English block overt movement, only attraction of the verb's phi features from V to AGR is licensed. By contrast, strong AGR features in French drive overt movement of the verb to AGR to check and erase its -Interpretable phi features.

According to the SBH only substanitve elements of the lexicon are represented in ECG, FCs are subject to a maturational schedule and do not come on-line until they have matured. A second major difference between ECG and adult grammar is that, while in the latter linguistic structures are networks of both thematic and non-thematic relations, in the former only thematic relations are realised because of the lack of FCs. In a typical child's utterance such as (7) where all the functional elements that are required in the adult grammar are missing, only thematic sisterhood relations inside the VP shell are represented:

(7) Her climbing ladder. (Jem, 2:0) (from Radford, 1990: 176)



Both nominals are non-target, the subject is in accusative rather than nominative case, and the object *ladder* is determinerless. Because a functional D category is not available at this stage, the two nominals are not subject to the requirements that apply to DPs, checking of the -Interpretable nominative Case feature by raising to spec-AGRP for the subject, and the overt marking of specific reference through a Determiner for the object DP. As for the verb form *climbing,* it contains the aspectual progressive marker *-ing*, although this is not considered to be a functional element, Aspect is initially encoded trough a process of lexical affixation rather than syntactic affixation (cf. Tsimpli, 1996). There is no aspectual auxiliary *be*, which is clearly required both for

Tense specification and for checking the Case feature of the subject in spec-AGRP. At the stage when Jem uttered (7) he was obviously not yet aware of any of the functional elements that are required in the adult grammar. What he does seem to have mastered though, is the argument structure of the verb *climb* which takes an agent subject and a patient object. Judging by the word order of the two arguments with respect to the verb, Jem must also have figured out that English is a head-initial language where complements are generated to the right of the verb, and that specifiers are to the left of the verbal head and hence precede it.[7]

Radford (1990, 1992) argues that early child grammars of English lack both T and AGR, in addition to other FCs such as D and C. Children are not aware of T and AGR features, they are not represented on the lexical items they draw from the lexicon for the Numeration, therefore the purely uninflected substantive elements appear in the derivation. Because they lack any formal features such as Case or phi features, the lexical items that appear in ECG of English are not subject to the checking requirements that govern the syntactic behaviour of lexical items that do carry formal features, hence no movement is necessary outside the VP. The starting point of Radford's argument is that the absence of T and AGR predicts that children will not produce verbal forms inflected for Tense and Agreement, if they did the operation Attract would drive percolation of the relevant weak features from the VP to T and AGR, and VP-internal subjects would have to raise overtly to spec-AGRP to check their nominative Case. However, these very operations are barred by the absence of target FCs.

At the core of Radford's line of argumentation is the central assumption of a direct relationship between overt morphological inflections and formal features. Inflectional morphology is taken to be the overt realisation of formal grammatical features. Take, for example, the lack of grammatical gender marking on nominal elements in English. Because there is no overt morphological evidence that the feature [gender] is relevant for the nominal system of English, this feature is not considered to be part of the inventory of formal features that are relevant in the grammar of the language. By contrast, in languages where grammatical gender is morphologically marked, it is taken to be the overt reflex of a formal feature that is part of the inventory of formal features of that language. By this rationale, in a given language only those

[7] The fact that parametric choices such as whether a language is head-initial or head-final appear very early in language acquisition, despite the existence of a number of other non-target properties of child language at this early stage, has led Wexler (1998) to propose that Very Early Parameter Setting is operative in child language.

features for which there is morphosyntactic evidence will be considered to be part of the grammar. Applying this claim to child language, a child will be credited with knowledge of only those formal features for which there is morphosyntactic evidence in his or her production. This conservative approach to the assessment of the structural complexity of child grammar is shared by most researchers advocating some version of the Weak Continuity Hypothesis, an approach that will be discussed in more detail in section 4.3.4.

The first piece of evidence that Radford adduces for a non-existent T node is the absence of modals from early child grammars of English. Modals are supposed to be base-generated in T in adult English, if T is initially missing one would not expect to find modals in these initial stages. Numerous studies (Brown, 1973; Hyams, 1986; Aldridge, 1989, among others) confirm Radford's predictions, modals are systematically absent in English ECG, both in naturalistic production and in elicited repetition tasks as shown by experimental studies by Brown and Fraser (1963), Brown and Bellugi (1964) and Ervin-Tripp (1964):

(8)	ADULT MODEL	CHILD'S IMITATION	
a.	Mr Miller *will* try	Miller try	(Susan 2;4)
b.	I *will* read the book	Read book	(Eve 2;1)
c.	I *can* see a cow	See cow	(Eve 2;1)

(from Radford, 1992: 27)

In Radford's model the child utterances in the repetition task are simple VPs where the children have maintained the verb-initial word order and the internal argument is minimally expressed as an indeterminate nominal, while the external argument is missing and is represented as a phonologically null NP (*np*). All other functional elements, including the nominative case-marked subjects and the modals, are missing as would be expected if T, and AGR, are not represented in the child's grammar.

Another obvious gap in early child grammars of English is the absence of Tense inflections represented by the regular affixation of *-ed*.[8]Traditional morpheme

[8] Irregular past tense forms cannot be taken as evidence of a productive T system unless it is clear that they are being used contrastively with the corresponding present tense forms, it could simply be that the child has learnt a past tense form but that she uses it as if it actually were a present tense form. This could of course be true also of *-ed* past forms, although because this is a productive affixation

studies (de Villiers and de Villiers, 1973; Brown, 1973) and a host of other studies on the acquisition of English (Bloom, 1971; Radford, 1990, Tsimpli, 1996) report that children systematically omit Tense marking when it is actually required by the adult grammar. Likewise, Agreement markings in the form of dummy auxiliary *do* , auxiliaries *have* and *be*, copula *be* and present tense verbs inflected for singular number (*-s*) are generally missing from early grammars of English.

Absence of Tense marking is also reported by another proponent of the SBH for Early Swedish.[9] In a study of three Swedish children between the ages of 1;8 to 2;3, Platzack (1990, 1992) clearly shows that, at least for the very initial period of acquisition (MLU range 1.19-1.52) there is no need to postulate the existence of a functional category outside the VP. At this stage 11 verbs out of a total of 36 are used with tensed forms, 4 of which are also used in the corresponding infinitival form. Platzack speculates that these tensed forms are no more than lexical variants of infinitives; they do in fact occur in the same position as the other 26 infinitivals occurs. However, although Swedish is a V2 language where V raising to a higher functional projection is required for all tensed verbs in matrix clauses, the V2 phenomenon is clearly observable only when a non-subject precedes the verb which in turn precedes the subject:

(9) [CPDett Csatt [IP en katt på mattan]].
"There sat a cat on the-mat"

When a subject is in spec-CP the word order is the same as if the verb had remained in the VP:[10]

(10) Klocka ligge där. (Freya 2;1)
"Watch lies there"

(from Platzack, 1992)

---

rule it would be easier to disambiguate whether the child is using it to mark past tense or whether it is an unspecified affix which indiscriminately marks any verb regardless of a past or present interpretation. As for the *-s* affix, its status has been variously described as being a Person, Number or Tense marker, Kayne (1989) suggests that *-s* in fact marks singular number. Here the *-s* marker will be treated as marking AGR rather than Tense.

[9] Agreement is non-existent in Swedish, Tense and Mood are however contrastively distinguished by different inflectional endings, usually the finite verb ends in *-r* (present tense), or *-de/-te/ø* (past tense), whereas the infinitive ends in *-a*.

[10] Although Swedish is a V2 language like German and Dutch, it is not head-final in the VP like German and Dutch which are OV languages, Swedish is a VO language.

The utterance in (10) is also compatible with an account that posits that the verb has actually raised out of VP to an intermediate functional projection. Platzack is nevertheless inclined to treat these early occurrences of tensed verbs as having remained in VP. He presents crucial evidence from 15 utterances where a non-subject is in first position, 11 of these contain an infinitive which, as expected, has clearly remained in VP. However 4 examples contain tensed verbs which should have raised out of the VP if an additional functional projection where present to check the [+finite] feature. This is not the case, the verb, even though it is appropriately inflected, remains in VP, as the word order lacking V2 shows:

(11) Julklapp Embla har.
"Christmas-present Embla has"

(from Platzack, 1992: 69)

For Platzack this is uncontroversial evidence that inflected and uninflected forms are not used contrastively by the children, they appear in free variation as lexical variants of the infinitive. This observation has serious implications for the status of inflected forms in early child language, one should be careful to link the presence of inflected forms with the existence of a FC. For a desirable principle of structural economy and parsimony, overt syntactic evidence or proof of productive and contrastive use of inflected forms should be required at all times to justify the postulation of a FC. Although Embla uses a number of inflected verbs, she does not always treat them as verbs carrying a [+finite] feature which must obligatorily be checked by raising the verb to an appropriate functional category outside the VP. Out of 15 examples of non-V2 verbs in Embla's corpus, there are 4 instances of non-V2 inflected verbs, even though these verbs are morphologically target-like, their syntactic behaviour betrays their non-target treatment as non-finite forms.

Platzack's observations highlight a non-trivial problem of data assessment in child language. It is necessary to evaluate a number of sources of information and to cross-check between them, in order to obtain a satisfactory degree of accuracy when making claims about the state of a child's grammar. The morphological information in Embla's case, presence of inflected verbs, must be verified against syntactic information, V2 vs. final position, before one can be sure that the inflected forms truly represent instances of target-like finite forms. Qualitative data are also important in deciding to what extent a child can be credited with mastery of formal features and corresponding FCs. If, for example, a child's performance on past tense forms were

95% correct, but it turned out at a closer inspection that out of 50 tokens only 2 verb types were represented, one would obviously have to judge this data a lot more cautiously than if the child had 15 or 20 verb types in his repertoire of past tense forms.

From this brief overview of children's performance on verbal inflectional morphology, one could venture the hypothesis that children have difficulties with the acquisition of inflectional affixes *per se*. At a closer inspection this does not seem to be the case; there is evidence of early productive use of the aspectual markers *-ing* and *-en*. At the same time as children do not master Tense marking, they are sensitive to Aspectual distinctions. The *-ing* aspectual affix marks both progressive imperfective aspect, while the past participial *-en* marks perfective aspect. Following Smith (1991), Tsimpli (1996) proposes that two independent aspectual components interact to give rise to the overall aspectual meaning of a given sentence: situation type and viewpoint. Substantive elements express the situation type, i.e. the choice of verb will determine what situation type is intended. Tsimpli lists five basic situation types: states, accomplishments, achievements, activities and semelfactives.

(12)a. know the answer (state)
b. stroll in the park (activity)
c. build a house (accomplishment)
d. reach the top (achievement)
e. cough (semelfactive)

An additional set of three binary features is assumed to interact with the aspectual interpretation of the above situation types: [± static], [± telic], and [± durative]. According to Smith (1991) the distinction between situation types derives from human perceptual and cognitive abilities, and is therefore supposed to hold universally. Clearly, however, there will be crosslinguistic variation as to how these situation types are actually grammaticalized in the various languages.

Viewpoint, by contrast, is encoded by morphological affixation and involves a tripartite distinction of perfective, imperfective and neutral.[11] The perfective specifies initial and final points, the imperfective specifies duration, and the neutral specifies the initial point and at least one internal stage. In English viewpoint is morphologically

[11] According to Smith (1991) the neutral viewpoint allows both imperfective and perfective readings, and Tsimpli (1996) argues that infinitival forms in child language should be regarded as having a neutral viewpoint, their interpretation being assigned on the basis of contextual information.

collapsed with Tense, thus Past Tense is also perfective, while Present Tense is imperfective, like *-ing* forms, but while Present Tense has a non-progressive reading, *-ing* forms are clearly progressive. Various researchers have proposed that children are sensitive to Aspect distinctions before Tense distinctions. Antinucci and Miller (1976) provide evidence that from very early on, before they contrastively mark Tense distinctions, Italian-speaking children are sensitive to the difference between stative and non-stative verbs, and use imperfective markers only with the latter and not the former. In Modern Greek, a language where Aspectual markers are distinct from Tense markers, both Katis (1984) and Tsimpli (1996) show that in a stage where children are not sensitive to Tense marking yet, Aspectual distinctions are already operative. Data from the acquisition of Serbo-Croatian, another language where Apect and Tense markers are distinct, indicates that it is indeed the case that Aspect emerges earlier than Tense (Raduloviç, 1975).

As far as English is concerned, there is abundant evidence that *-ing* forms and past participial *-en* forms are used early on in acquisition, in fact Brown (1973) reports that the *-ing* morpheme is one of the first to be mastered by Adam, Eve and Sarah. Perfectivity is expressed early on through the use of perfective *-en* forms that are correctly used only for non-stative situation type verbs (typically accomplishment and achievement predicates). *-Ing* forms, as well as infinitivals, express the imperfective reading with the progressive reading available for the former but not the latter.[12] The crucial factor for the SBH is that the encoding of Aspect is a purely lexical operation that does not involve formal feature checking in the domain of a FC. The emergence of aspectual markers is thus not problematic for the SBH and is in fact stronger evidence for the fact that the initial omission of functional affixation expressing Tense and Agreement cannot simply be explained away by difficulties in morphological learning, but it truly reflects the unavailability of FCs.[13]

Summing up, proponents of the SBH argue that representations of children's early utterances need not include the presence of FCs outside the VP. Drawing mainly from English and Swedish child language, several pieces of evidence have been

---

[12]Hyams and Sano (1994) and Hyams (1996) also treat children's early bare *-en* participial forms, and also the earliest occurrences of the *-s* marker, as predicates expressing aspectual rather than temporal information.

[13] It could be argued that aspectual markers do not involve the mastery of a paradigm. Unlike Agr and possibly Tense affixes, they apply directly to a verbal stem. The relevant piece of knowledge entails mastery of the difference between stative and non-stative verbs, more of a semantic notion than a purely grammatical one.

presented in support of this hypothesis: absence of modals, auxiliaries *be* and *have,* missing Tense and Agreement inflections, the presence of non-V2 tensed verbs in the grammar of Early Swedish, and finally the presence of lexically represented Aspect.

In the next section I will turn to consider two alternative accounts of the initial status of FCs in ECG: the Strong Continuity Hypothesis and the Weak Continuity Hypothesis.

4.3.2. The Strong Continuity Hypothesis

The core assumption underlying the SBH is that FCs are not initially part of the child's grammar and that they are subject to a maturational schedule. Some researchers have criticised this view of language acquisition on the grounds that it poses a discontinuity problem with the adult grammar. As an alternative to the SBH two parallel proposals have emerged in the literature: the Strong Continuity Hypothesis (SCH) and the Weak Continuity Hypothesis (WCH). I will start by outlining the basic tenets of the SCH together with the empirical evidence that researchers working in languages as diverse as English, French, German and Italian have offered in support of their hypothesis.

In order to minimize discontinuity with adult grammar the null hypothesis is that children's access to FCs is the same as adults (Pinker, 1984; Poeppel & Wexler, 1993; Deprez & Pierce, 1993, 1994; Hyams, 1986, 1987, 1992, 1994, 1996; Pierce, 1992; Verrips & Weissenborn, 1992; Hoekstra & Hyams, 1998). By this rationale, both UG principles and full knowledge of the inventory of FCs are available to children from the earliest stages of acquistion. The assumptions underlying the SBH are challenged both on theoretical and empirical grounds. From a theoretical point of view, Hyams (1992: 371) questions the very nature of the "learning and/or maturational mechanisms responsible for the emergence of these [functional] categories". In the absence of an adequate explanatory model of how FCs are supposed to emerge out of nothing, Hyams, together with other advocates of the SCH, proposes that the only viable alternative is to assume that ECG differs minimally from adult grammar and that it does indeed incorporate FCs.

A second objection to the feasibility of the SBH comes from an accumulating body of crosslinguistic evidence which indicates that the existence of FCs in ECG is justified by the presence of productive use of Tense and Agreement morphology, the

presence of clitics, correct nominative Case assignment to subjects of tensed predicates, correct placement of tensed verbs with respect to negative markers, and matrix V2 phenomena in languages like German and Dutch.

For Italian, Hyams (1986) and Guasti (1993/94) report that there is evidence of productive use of Agreement morphology, at least in the singular paradigm.[14] On the basis of distributional evidence in the use of tensed verbs and infinitivals, Guasti argues that, as early as 1;8, the three monolingual children in her corpus (age range 1;8-2;7) can distinguish between appropriate syntactic contexts for finite and non-finite verbs. The overwhelming majority of infinitives appear as the complements of tensed verbs or governed by a preposition:

(13)a. Voglio bere. (Martina, 1;10)
"I want to drink"

b. Posso entrare? (Diana, 2;1)
"Can I come in?"

c. pe' c(u)ocere. (Martina, 1;8)
"to cook"

d. per lavale i piatti. (Diana, 2;1)
"to wash the dishes"

(from Guasti, 1993/94: 4)

Although the Italian children in this study are aware of the distinction between finite and non-finite forms, nevertheless they also produce a number of non-target Root Infinitives (RIs), that is infinitives in matrix clauses where a tensed form is required in the target grammar. Guasti points out though, that the percentage of RIs is lower in Italian when compared to other languages, such as Swedish, German and French, where this phenomenon is known to occurr. However, when the number of bare past participles that also occur with some frequency at this stage is added to the count of non-target non-finite forms, the figure reaches comparable levels to those in other languages with RIs. Guasti tentatively speculates that the lower percentage of RIs in Italian child language could be related to the strength of the AGR node and the fact that generalized verb movement to the highest inflectional node is required for all verbs in Italian (cf. Belletti, 1990).[15]

[14] However see Pizzuto & Caselli (1992) for an alternative evaluation of Italian children's mastery of the inflectional paradigm.

[15] See Wexler (1998) for an in-depth treatment of RIs in child language and the connection between null subejcts and Optional Infinitives.

The distribution of clitics in early child Italian also attests to the operative distinction between finite forms , which will have raised out of VP to check their phi features in AGR, and infinitivals that need not raise to a higher functional position. In adult Italian clitics are placed to the left of the tensed verb and to the right of the infinitive[16]:

(14)a. Lo voglio.
"I want it"
b. Voglio mangiarlo.
"I want to eat it"

Although the number of clitics overall is low for all three children in Guasti's study, they are always correctly placed.

Not only do Italian-speaking children distinguish between finite and non-finite forms, they also show early productive knowledge of the Agreement paradigm as Guasti (1993/94: 22) states: "children do not make systematic agreement errors. All singular person inflections emerge earlier than the plural ones and are mostly used correctly". Hyams (1986) also reports that her findings confirm that Italian children have mastered the present tense paradigm by roughly age 2. There are no specific claims as to the relevance of Tense distinctions in the grammar of Italian-speaking children, and although there is evidence that past participles are frequent from the beginning, they are initially produced as bare participles without the accompanying auxiliary that is required in the adult grammar. These bare participles have been variously analysed as adjectival forms (Volterra, 1976), verbal forms marked for Aspect but not Tense (Antinucci & Miller, 1975) or instances of Root Infinitives (Lyon, 1997).

The available data on the acquisition of Italian indicate that there is reason to assume that at a relatively early age children have already a notion of finiteness which governs verb movement and clitic placement, and that at the same time they display considerable knowledge of the Agreeement paradigm. Taken together these pieces of evidence point towards the availability of a functional Agreement category to which inflected verbs raise in order to check and erase their strong -Interpretable phi features.

---

[16] An alternative to (14)b is (1) below where the clitic has climbed to the left of the tensed verb:

(1) Lo voglio mangiare.

Studies on the acquisition of French negation (Pierce; Déprez & Pierce, 1993) also show that children are sensitive to the finite/non-finite distinction and that only tensed verbs raise to the left of the negative marker *pas*, while infinitival forms (both infinitives and bare participles) remain in their base-generated position to the right of Negation:[17]

(15)a. Pas manger la poupée.[18]
"Not eat the doll"
b. Veux pas lolo.
"Want not water"

(From Deprez & Pierce, 1993: 40)

Data on the distribution of subject pronouns in early child French substantiates even further Déprez & Pierce's argument for the availability of an Inflectional projection in the children's grammar. French subject clitics are assumed to be clitics which raise to I in the adult language, it is therefore significant that subject pronouns in French child language are only found in combination with finite verb forms, the use of subject clitics with non-raised infinitives is unattested:

(16)a. Il est pas là. (Nathalie, 2;2)
"He is not there"
b. Et je veux. (Nathalie, 2;2)
"And I want"
c. Elle dort. (Daniel, 1;8)
"She sleeps"

Additional evidence for the differential syntactic behaviour of finite vs. non-finite forms is provided by the study of Andreas, a 25-month-old German child, in Poeppel & Wexler (1993). The criterion used in this study to establish finiteness classes all verbs ending in *-en*, the canonical infinitival marker, as non-finite; all other forms are considered to be finite. Of course *-en* in adult German is also the marker for

[17] French clause structure is assumed to include a NEGP to the left of the VP headed by the clitic head *ne* and the specifier *pas*. The head *ne* is assumed to cliticize onto the tensed verb, hence any diagnostic for verb movement relies on the position of the verb with respect to the negative marker *pas* which is fixed in spec-NEGP position.

[18] Déprez & Pierce (1993, 1994) account for postverbal subjects with non-finite verbs in the grammars of young French-speaking children by proposing that spec-VP can either be to the left or to the right of V'. An alternative is that these postverbal subjects are instances of right-dislocation, a common occurrence in the adult grammar.

first and third person plural subject agreement, and in these circumstances it should obviously count as a finite ending. However, there is no indication in Andreas' production that at this stage he has analysed this affix as a marker of first and/or third person plural. In fact the few first and third plural subjects that do appear in the corpus, 11 altogether, are used with the corresponding singular agreement. In so-called V2 languages like German, Dutch and Scandinavian languages in general, finite verbs must appear in second position in declarative matrix clauses, this position is typically identified with C (cf. Haider & Prinzhorn, 1986; den Besten, 1989). C is the FC the finite verb must raise to to check its [+ finite] feature, therefore availability of V2 in German child language is taken as an indication that a C category, or minimally a F category expressing finiteness , is present in the child grammar (cf. Clahsen & Penke, 1992).

Out of a total of 251 utterances with three or more constituents in the Andreas corpus, 197 finite forms appear in V2 position, vs. only 11 in sentence-final position; while only 6 non-finite forms are found in V2 and 37 are in sentence final position. Poeppel & Wexler conclude that finiteness is a relevant notion in Andreas' grammar and that V is in C position.

Another case study of German child language is reported in Verrips & Weissenborn (1992). The informant is Simone, a German-speaking child studied from 1;9.11 to 2;2.21. In line with the results of Poeppel & Wexler's study, here too the findings confirm that from the earliest multiword constructions Simone can distinguish between finite and non-finite verbs: the overwhelming majority of verbs in V2 position are finite, while sentence-final position is reserved to non-finite forms. Contra de Haan (1986) for Dutch, Verrips & Weissenborn argue that there is no evidence in the Simone corpus that the V2 position is lexically restricted to modals and auxiliaries; they also dispute Jordens' (1990) claim that only activity verbs are found in V2 in early Dutch child language. Clahsen & Penke's (1992) study on the same Simone corpus reports that the child's V2 verbs are initially only those marked by *-t*, which they claim to be an intranisitve marker or possibly some sort of aspectual marker of accomplishment rather than an agreement marker for third person singular. Verrips & Weissenborn defend their hypothesis of generalized V2 movement for finite forms driven by the syntactic requirement to check the [+finite] feature in C. They cannot find any evidence of lexical restrictions either for V2 or for the elements which appear in first position, spec-CP in their analysis. If the finite functional projection hosting finite verbs in V2 position is indeed a CP, its specifier position should be available to

subjects, objects, locative phrases, adverbials and wh-prhases alike. The results confirm that the initial position is not exclusively reserved for subjects, out ot 417 utterances containing a topicalized element, 130 are non-subjects (28%). Verrips & Weissenborn therefore conclude that movement to V2 of finite forms is clearly attested well before there is substantial evidence for the mastery of the Agreement paradigm. Their initial hypothesis on the independence of syntactic movement to C from the acquisition of Agreement inflection is thus confirmed. Contra Clahsen & Penke (1992), they state that there are no "syntactic consequences for verb placement observable in the data of the children that uniquely depend on the presence of agreement" (Verrips & Weissenborn, 1992: 320).

A number of crosslinguistic studies in languages where there are clearer morphosyntactic cues for the presence of FCs than there are in English suggest that it is necessary to assume that children have access to FCs from the earliest stages of acquisition. What is more, although the empirical evidence alone confirms that ECG cannot be universally a lexical only grammar, proponents of the SCH argue that the availability of FCs is a theoretically necessary postulate of language acquisition. In the absence of a convincing explanation of how FCs could possibly emerge and/or mature, the only viable alternative is to assume that they are always part of the child's grammar, even when there may be no clear morphosyntactic correlates present in the data. Hyams (1992: 390) does not believe that in order to assume that a given FC is present one should necessarily rely on the presence vs. absence of particular lexical items:

> "... the small clause hypothesis also suffers from a particular conceptual problem, which is that it rests on the dubious assumption that one can argue from the absence of particular lexical items in the child's language to the absence of a grammatical category in his/her grammar"

Nevertheless, despite the empirical evidence indicating that in a variety of languages there is reason to believe that a VP analysis including only lexical categories could not do justice to the data, there are also clear indications that children's grammar are not target grammars from the very beginning. It has been noted time and time again that, in the initial stages of acquisition, syntactic phenomena which are obligatory in adult grammar are optional in child grammar. Two extensively studied areas of early child language are null subjects and root infinitives. It is a well-attested fact that even children acquiring non-null-subject languages initially tend to optionally omit subjects which are obligatory in the adult grammar. Hyams (1986) originally proposed a parameter missetting analysis for this phenomenon in child English and proposed that English-speaking children had erroneously set the [pro-drop parameter] to the [+]

value, for all intents and purposes the children treated English as if it were Italian, producing *pro* subjects. This original proposal has since been subjected to extensive reformulation due to both theoretical and empirical inadequacies. From a theoretical point of view there are difficulties with the requirements imposed by the identification conditions on the supposed *pro* subjects in child English. Identification of definite reference as that intended for null subjects in child English is warranted by a rich Agreement system (cf. Rizzi, 1986) which is not the case for English. Secondly, null subjects in child English do not have the same distribution as null subjects in adult Italian, there is no evidence that English speaking children drop subjects in embedded clauses, an option which is perfectly grammatical in a true pro-drop grammar. Valian (1991) also demonstrated how the proportion of null subjects is in fact very different between the language of children who are acquiring a real pro-drop language like Italian, and the language of children that are acquiring a non-pro-drop language like (American) English. From age 1;6 to 2;6 the Italian-speaking children in Valian's study include subjects in a minority of their sentence containing a verb (30%). The group of American children participating in the study produced overt subjects more than twice as often (70% in the lowest MLU group, MLU 1.53-1.99), with an even higher proportion in the group with MLU ranging from 2.25 to 2.76: 84%.

In a reanalysis of the null subject phenomenon, Hyams (1994, 1996) and Hyams & Sano (1994) have proposed a new explanation for English children's null subjects in terms of underspecification of the I system. Starting from the empirical observation that children may optionally specify a verb as [-finite] in a context where the adult grammar obligatorily requires that it be [+finite], Hyams & Sano (1994) and Hyams (1996) suggest that, although the I system is part of the child's grammar, at this stage it may be qualitatively different from the adult I system in that it is underspecifed. While finiteness must be morphologically realised in the target grammar to specify a verb as [+finite], in the child grammar temporal specificity may optionally remain unmarked. A [-finite] I will therefore license a PRO subject whose null Case is appropriately checked by a minimal I, that is an I devoid of lexical content. In this new analysis, English null subjects are the consequence of the underspecification of the I node, and not, as previously argued, *pro* subjects resulting from the missetting of the pro-drop parameter.

More recently Hoekstra & Hyams (1995, 1998) have linked the presence of Optional Infinitives (OIs) to the underspecification of the feature Number. If Number is unspecified in a language where the only marker for finiteness is number

morphology, then there is no morphological device to distinguish between finite and non-finite forms, hence the presence of RIs. Hoekstra & Hyams (1996) divide languages into a-type languages and b-type languages according to whether there are Person distinctions in the inflectional paradigm (a-type languages) or not (b-type languages).

English is a b-type language, *-s* simply marks a singular Number distinction, while plural Number is marked by the bare form. Again no Person or Tense distinctions. By contrast, in languages like Italian, Spanish, and Catalan, where the inflectional paradigm marks Person distinctions, each person has a distinct marker:

(18) Italian *parlare* ("to speak")

io parl<u>o</u>	noi parl<u>iamo</u>
tu parl<u>i</u>	voi parl<u>ate</u>
lui parl<u>a</u>	loro parl<u>ano</u>

French, a language in which there is evidence of RIs, has a mixed system. The paradigm marks person distinctions in the plural, but not in the singular:

(19) French *parler* ("to speak")

[j parl]	[nu parl]
[tu parl]	[vu parl<u>e</u>]
[il parl]	[il parl]

German, another language where Person distinctions are marked not only in the plural but also in the singular, also exhibits the RI phenomenon in child language:

(20) German *spazieren* ("to walk")

ich spazier(<u>e</u>)	wir spazier<u>en</u>
du spazier<u>st</u>	ihr spazier<u>t</u>
er/sie spazier<u>t</u>	sie spazier<u>en</u>

Given the presence of Person distinctions in the French plural, and both in the singular and the plural for German, the predictions made by the underspecification of

Number hypothesis are that RIs should not surface in French and German child language. Hoekstra & Hyams (1996) suggest that in order to understand the behaviour of RIs in French and German, one must take into account what they call the *avoid plural phenomenon.* It is a well-know fact that initially children tend not to refer to plural entities and therefore they overlook the plural half of the paradigm.[19] If children only focus on the singular paradigm, then French is a language where there are no Person distinctions and it is therefore treated as such by the children.

For German the explanation is somewhat more complicated. At the stage when German children produce RIs, not only do they not use plural inflections, but they also have not yet acquired the 2 p.s. marker *-st* (cf. Clahsen & Penke, 1992). Under these circumstances, German looks very much like Dutch where the 1st person is unmarked and the *-t* marker specifies Number and not Person. If ECG has the option of leaving Number unspecified, and no Person marking in child German is yet relevant, then the presence of RIs is accounted for.

Although appealing in principle, Hoekstra & Hyams' proposal suffers from a number of technical problems. Firstly, the authors classify languages according to whether they mark Number, as in the case of Dutch and English, or whether they mark Person (Italian, Spanish and Catalan), or Tense (Japanese). The classification system seems to be working well with clear-cut cases such as the ones listed above, but it encounters some difficulties with French and German, where Person distinctions are indeed marked, although not as consistently as in prototypical Person marking languages like Italian, Spanish, and Catalan. French has three different markers: a ø marker for the three singular persons and for 3 p.p., a 1 p.p. marker [o], and a 2 p.p. marker [e]. Therefore, although there is no clear Person marking for the singular persons, Person distinctions are clearly available in the plural and from a purely taxonomic point of view French should be considered an a-type language.

The German case is even more complicated, since Person distinctions are marked both in the singular and in the plural, cf. the paradigm in (20) above, hence German too should be classified as an a-type language. If this were the case Hoekstra & Hyams' proposed correlation between RIs and purely Number marking languages would not hold, since RIs are also found in French and German which do display Person marking, albeit to a lesser degree than Italian, Spanish, and Catalan. This is

[19] See Pizzuto & Caselli (1992) and Guasti (1993/94) for relevant data on Italian, Rubino & Pine (1998) for Brazilian Portuguese, Brown (1973) for English, and Clahsen, Penke & Parodi (1994) for German.

where the second major problem with the approach lies: in order to reconcile the presence of RIs in French and German child language with the existence of Person markings in these two languages, Hoekstra & Hyams have to propose that children do not in fact pay attention to these inflectional markings at all. Plural inflections are ignored on the basis of some general avoidance strategy. As for singular Person markings in German, it is a well-known fact that 2 p.s.*-st* is late to be acquired, and in the absence of *-st* and the unmarkedness of 1 p.s. *-e*, the 3 p.s. *-t* inflection could well be interpreted as a Number marker rather than a Person marker. Assuming that this may be a possible scenario, what does it actually mean to say that children ignore these inflections? Although it is true that they are late in producing them, they cannot ignore them in comprehension, and if they do hear them and can interpret them correctly, this will surely have some impact in the child's representation of the inflectional paradigm. At least from the comprehension side of things, there must be some awareness on the part of the child of a fairly rich and diverse inflectional paradigm in French and German, as opposed to the uniformity of English and Dutch where there really is only an opposition between singular and plural forms.

Despite the difficulties that arise from the application of Hoektra & Hyams' underspecification of Number hypothesis to languages like French and German, the underlying insight clearly captures a fundamental typological distinction in the distribution of RIs in child language. A tentative solution proposed here is to link the presence of RIs to the morphological uniformity of the inflectional paradigm, thus sidestepping the issue of Person and Tense marking vs. Number marking. The basic idea is derived from the notion of morphological uniformity proposed by Jaeggli & Safir (1989:30):

> a. "An inflectional paradigm P in a language L is morphologically uniform iff P has either only underived inflectional forms or only derived inflectional forms."
> b. "Null subjects are permitted in all and only those languages with morphologically uniform inflectional patterns."

It can be no coincidence that RIs do not appear in those languages which allow referential null subjects, like Italian. Because this is the case, it is highly probable that the same property that licenses referential *pro* subjects also disallows RIs. Hamann (1996) also capitalises on the notion of morphological uniformity to account for a puzzling phenomenon of null subjects in 3-year-old German-speaking children which look very much like referential *pro* subjects and not simply instances of topic drop. She distinguishes between *formal morphological uniformity* of the Agreement system

for which it is sufficient to have a slot for each inflection, i.e. an ending for each person, although they do not all have to be different (the case of German); and *functional richness* of the Agreement system, whereby AGR is marked as [+pronominal] and at least five of the six inflectional endings are different (the case of Italian). In a language where AGR is functionally rich, referential null subjects are allowed and RIs are virtually unattested. Italian, Spanish and Catalan all have a functionally rich AGR, and six different Person markers. By contrast, French and German do not have a functionally rich AGR, Person markings do exist (unlike in English and Dutch), but they do not meet the criterion of at least five different Person endings out of six. In languages with a functionally rich AGR, finiteness is always realised by a unique ending marking Person, therefore even if Number may be underspecified there are no observable non-finite effects resulting in RIs.[20] In languages where Person marking does exist, but is not so consistent as to qualify the language's AGR as functionally rich, if Number is underspecified there may be no other way for the language to mark finiteness, hence the appearance of RIs.

Further support for the underspecification of Number hypothesis comes from a third type of language: Japanese. Hoekstra & Hyams define Japanese as a c-type language where only Tense is morphologically marked. On this assumption, similarly to the pro-drop languages situation where RIs are blocked by the specification of Person, RIs in Japanese should not occur because, although Number may be underspecified, Tense morphology marks the verb form as [+finite]. The prediction is confirmed by a study of five Japanese children (ages 1;11 to 2;11) where Sano (1995) reports the total absence of root non-finite forms. The Japanese data is particularly significant for Hoekstra & Hyams' hypothesis, in the light of proposals such as Wexler's (1994, 1998), where the underspecification of Tense is assumed to be responsible for the RI phenomenon. Contrary to the evidence, a language like Japanese would be predicted to show RI effects by Wexler's underpecified Tense hypothesis.[21]

---

[20] See Lyon (1997) for the proposal that the underspecification of Number has observable morphological consequences for participial structures involving unaccusative verbs in Italian child language.

[21] Hoekstra & Hyams (1996) also note that Scandinavian languages should be predicted to pattern with Japanese rather than with the other Germanic languages. Yet they show a significant RI effect. The solution is to treat the uniform *-er* marker in the present tense paradigm not as a present Tense marker, but as a degenerate Number marker. If *-er* was a Tense marker then the disappearance of *-er* in the present tense passive vs. the presence of the past marker *-te* in the past tense passive would remain unexplained:

(i)	jeg valger	jeg valges
	"I choose"	"I am chosen"
(ii)	jeg valgte	jeg valgtes

Because RIs are non-finite forms where no Tense chain has been formed, the temporal anchoring cannot take place through syntax, but the fact still remains that children's RIs must get an interpretation. The proposal is that when the discourse anchoring is unavailable via syntactic devices it takes place through pragmatics. Children assign a default (usually) present Tense interpretation to their RIs by exploiting temporal coreference with a discourse antecedent located in their speech time. At the OI stage children's grammars are not as constrained as adults, although RIs are possible in the adult language too, they are subject to narrower pragmatics constraints than children's RIs. As noted also by Chien & Wexler (1990) for the nominal binding domain, and Schaeffer (1997) for object scrambling in Dutch and object clitic placement in Italian, there seems to be a dissociation between the syntactic and the pragmatic modules, it is only when children's pragmatics reaches the adult target that these ungrammatical pragmatic options are discarded.

Ultimately the SCH achieves both goals of descriptive adequacy and theoretical explanation. Firstly, it gives a principled account for the empirical consideration that from very early on child grammar is not as deficient as claimed by proponents of the SBH. Crosslinguistic empirical evidence requires the assumption that FCs are part and parcel of ECG: early knowledge about verb movement, the distribution of null subjects, and sensitivity to inflectional property of the language have all been presented in support of the presence of FCs.[22]Secondly, the Underspecification Hypothesis, explains why it should be that children's grammars may differ from the adult target even though they include FCs. The availability of a FC does not automatically imply that it will be qualitatively the same as the corresponding FC in the target grammar. In fact it is very likely that, to begin with, the child's FCs will differ in that they will be underspecified with respect to some features that are realised in the adult grammar. The locus of the underspecification is currently being posited in the interface between syntax and pragmatics by a number of researchers. If this turns out to be the case, the gap between child and adult grammar will have to be explained not so much as a deficit in the "syntax proper" understood as the result of the operations of the Computational system, but rather as a deficit at the Conceptual-Interpretive interface.

For all its merits, the SCH makes too categorical a claim about the initial status of children's mental representations. The central argument of the hypothesis is the *a*

"I chose" "I was chosen"

[22] In a slightly different model, Wexler (1998) has proposed that Very Early Parameter Setting (VEPS) and Very Early Knowledge of Inflection (VEKI) are assumed to be operative in ECG.

priori availability of FCs in ECG, although some researchers concede that FCs may be underspecified in the early stages of acquisition. The reason which is typically proposed for this very strong claim is that the assumption of existence of FCs in child grammar eliminates the discontinuity problem with the adult target. In this respect the SCH is supposedly more parsimonious than other developmental accounts that view FCs as emerging either from the maturation of grammatical principles, or from the lexical learning of language-specific features. Proponents of the SCH have often criticised these gradual development hypotheses for not providing a well enough defined answer to exactly *how* FCs are supposed to emerge in the absence of previous specification of what it is that the child has to converge on. For example, in response to the Weak Continuity argument that the acquisition of FCs is triggered by the acquisition of the corresponding functional morphemes, i.e. that syntactic acquisition is driven by morphophonological acquisition, Borer & Rohrbacher (1997: 26) maintain that "[i]n the absence of a pre-existing notion of grammaticalised tense, it is hard to see why the child would be driven to segment a phonological string into a lexical stem and a discrete tense morpheme". Borer & Rohrbacher's objection arises from a fundamental misunderstanding of Weak Continuity positions. It is not the case that researchers working from this perspective advocate a *tabula rasa* approach to language acquisition, whereby FCs are solely instantiated on the basis of morphophonological input. On the contrary, FCs are assumed to be available in principle to the child from the onset of acquisition, however their status in child grammar is not the same as in the adult grammar, there is a qualitative difference between child and adult language, and although the *kind* of knowledge is in principle the same, the extent to which this knowledge is incorporated differs in significant ways.

The differences between Strong and Weak Continuity approaches originate in the way in which the two hypotheses view the relationship between FCs and their overt realisation through morphophonological markings. The tendency in SC approaches has often been to separate the two issues of abstract mental representation of a FC, thought to be part and parcel of the child's grammar throughout development, and lexical knowledge of the corresponding language-specific morphophonological morphemes, which does not necessarily go hand in hand with the availability of a target mental representation of the FC. In other words, while WC perspectives view the relationship between the availability of FCs and their overt morphophonological realisation as an implicational one, if morpheme then FC, SC approaches have tended to dissociate the two, the existence of a FC is independent from any overt morphophonological realisation, therefore the absence of relevant functional morphemes is not compelling

evidence for the absence of that category. However, it is not entirely clear to what extent is it really legitimate to separate morphology from syntax in this way. A number of studies on the acquisition of German have shown a clear correlation between agreement-marked forms and V2 position on one side, and non-finite forms and sentence-final position on the other (Clahsen & Penke, 1992; Clahsen, Penke & Parodi, 1993/94; Clahsen, 1996). Similar studies in French and Italian have also found a correlation between finite forms and V movement to higher functional projections (Déprez & Pierce, 1993, 1994; Pierce, 1992; Guasti 1993/94). There does seem to be a one-to-one correspondence between the emergence of verbal morphology and V movement out of the VP to FCs in those languages with strong Agreement and Tense features.

On this subject, Borer & Rohrbacher (1997) propose that it is the very absence of morphological markings in ECG that gives credence to the idea that FCs are indeed present even when language-specific morphophonological markings have yet to be acquired. Their focus is on AGRS, but the implication is that the same arguments apply to other FCs. The basic idea is that if AGRS were truly absent from English ECG, one would expect multiple commission errors involving the random use of agreement markers.[23] The presence of AGRS, on the contrary, requires that the fully inflected lexical items that are drawn from the lexicon during the Numeration and are inserted into VP raise to check their phi features. In order for the derivation to converge, the nominal phi features carried by the verb in AGRS must match the nominal phi features carried by the subject DP in spec-AGRSP. If a feature mismatch occurs then the derivation crashes and the sentence is ungrammatical. This is the case when either Person or Number specifications, or both, on the subject DP do not much the Person and/or Number specifications of the verb, e.g. a 3 p.s. subject and a 3 p.p. verb. On the assumption that AGRS is always part of the child's grammar, the only way to prevent

[23]Pine & Martindale (1996: 374)) argue that, the interpretation of low error rates in child language as an indicator of grammatical competence, is particularly problematic "in the absence of any clear idea about the number of errors which one might expect". The authors propose a number of reasons why errors might be low which do not imply that children possess any productive syntactic knowledge. Firstly, the number of errors that may actually occur depends on the number of tokens produced. A child with a small number of rote-learned utterances would be 100% correct, but one would be clearly unwilling to credit her with productive syntactic knowledge. Secondly, many potential errors are unlikely to occur for various reasons which may not actually have anything to do with the child's grammatical competence, but may interface with phonotactic or pragmatic constraints. Although in principle a child might be expected to produce a number of possible errors, some are more possible than others. In the case of determiner combinations for example, "kick the my ball", where both a definite article and a possessive coexist, is more likely to occur than "kick the a ball" where both a definite and an indefinite article appear in the same DP. While definiteness and possession are not two mutually exclusive semantic notions (and in fat there are a number of languages where the equivalent of "kick the my ball" is perfectly grammatical); definiteness and indefiniteness are complementary concepts, and hence the sequence "the + a" is rather unlikely to be attested.

the derivation from crashing, in the absence of relevant morphophonological knowledge, is to supply a minimal non-agreeing form: the bare stem in English, the infinitive in German, and a participle in French, Greek and Italian. Failure to provide the correct agreement-marked form results in omission errors, rather than comission errors, whenever agreement morphology *is* used it is used appropriately. Borer & Rohrbacher (1997) present data from three German-speaking children on their correct use of present tense affixes, their Table 3 is reproduced below as Table I:

	Andreas (2;1)	Katrin (1;5)	Nicole (1;8)
1 p.s. *-e*	21/22 (95%)	-	-
2 p.s. *-st*	8/8 (100%)	11/11 (100%)	-
3 p.s. *-t*	22/23 (96%)	25/25 (100%)	2/2 (100%)

Table I. Correct use of present tense affixes in child German
(from Borer & Rohrbacher, 1997: 29)

Borer & Rohrbacher maintain that the data in Table I are clear evidence that children do use agreement morphology correctly in an overwhelming majority of cases as soon as agreement markers start being used. This is presumably due to the fact that children have an AGRS category in their grammar that requires either phi feature matching between subject DPs and verbs, best case scenario; or, failing that, it minimally requires that there should be no feature mismatching, the phi features on the subject DPs must not clash with those of a verb specified for different phi features, worst case scenario, where the verb is simply not specified for phi features (bare stem, infinitival form).

The evidence presented by Borer & Rohrbacher at best indicates that children can reproduce agreement patterns to which they are exposed in the input. It does not provide any conclusive evidence that children have productive knowledge of subject-verb agreement as would be required by the presence of a functional AGRS category. Firstly, only singular Number is represented in the data in Table I, and only one of the children has a sizeable number of tokens for all three Persons, one of the children only has two tokens of 3 p.s. Secondly, and most importantly, there is no qualitative analysis of the number of verb types with which these inflections appear, and no indication of what structures they appear in. Especially when the number of tokens is very low, it becomes particularly difficult to claim with any degree of certainty that the correct utterances in the child's production do indeed show that a FC requiring feature checking is actually projected. For instance, the significance of 100% correct agreement

in 8 cases out of 8 for Andreas' 2 p.s. contexts clearly depends on the type/token ratio. Should it turn out that one verb type alone accounts for 90% of the correct 2 p.s. verb tokens, the figure of 100% would obviously have to be reconsidered in the light of this information. Without stringent quantitative and qualitative criteria of productivity, low error rates alone do not provide conclusive evidence for the existence of FCs. It is precisely for this reason that researchers like Pizzuto & Caselli (1992) adopt very stringent measures of productivity in their analysis of the acquisition of Italian morphology. In their study, the point of acquisition of each morpheme is defined, following Cazden (1968: 435), as "the first out of three consecutive speech samples such that in all three a given morpheme is supplied correctly in at least 90% of the contexts in which it is clearly required". It is a well-known fact that in spontaneous child data number of contexts for a given morpheme may vary considerably from one recording to the next. In order to account for possible variablity in the number of obligatory contexts, Pizzuto & Caselli adopt de Villiers & de Villiers' (1973) methodological solution of including in the scoring of a given morpheme only those transcripts where at least five obligatory contexts for that morphemes are identified. From a more qualitative perspective, the criterion of productivity for verbal inflections also requires that a given affix be used with at least two different verbs in the same recording. In the absence of this qualitative information about number of verb types with which verbal inflections appear it is impossible to draw any definite conclusions about the status of FCs.

Moreover, contrary to what Borer & Rohrbacher argue, there is evidence that children do commit agreement errors where phi features of the subject and the verb do not match. Ingham (1998) presents data from Sophie, an English-speaking child (2;6-2;9) who, in addition to omitting the *-s* Number marker on lexical verbs an average of 94.4% of the time, also makes a number of comission errors in the use of the copula *be,* progressive *be*, and auxiliary *have.* There is no indication that Sophie uses the various inflected forms appropriately; Table II shows the child's distribution of *is/are* with the various persons:

Participants	*is*	*are*
3 p.s.	21	6
3 p.p.	4	6
1 p.s.	2	1
2 p.s./2 p.p.	6	-

Table II. Use of is/are (2;6-2;9)
(from Ingham, 1998: 64)

The data in Table II clearly show that *is* is not restricted to 3 p.s. contexts, more than a third of the time (37% of obligatory contexts) *is* appears in contexts other than 3 p.s. (12% in 3 p.p. contexts, 6% in 1 p.s. contexts, and 18% in 2 p.s. or 2 p.p. contexts). Similarly, *are* correctly appears in a 3 p.p. context only 46% of the time, the rest of tokens are with 3 p.s. or 1 p.s. subjects. Ingham does not provide quantitative data on auxiliary *have*, but he claims that the same semi-random use applies.

Rubino & Pine's (1998) study of subject-verb agreement in a child acquiring Brazilian Portuguese also reports a number of agreement errors. Table III below replicates Rubino & Pine's Table 5 on the proportion of incorrect verbal productions over the total of correct plus incorrect productions for each different person:

	Agreement errors	N of occurrences	% Incorrect
1 p.s.	25	297	8.4
2 p.s.	0	188	0
3 p.s.	5	929	0.5
1 p.p. A Gente	4	17	23.5
1 p.p. Nos	0	10	0
2 p.p.	-	-	-
3 p.p.	10	23	43.5

Table III. Proportion of incorrect productions over total number of productions
(from Rubino & Pine, 1998: 44)

The largest proportion of errors is in 3 p.p. contexts, followed by 1 p.p. *A Gente* contexts, and 1 p.s. contexts. All the errors in 1 p.s. contexts involved the inappropriate use of a 2 p.s./3 p.s. verb inflection (2 p.s. and 3 p.s. verb inflections are homonymous forms). In a number of cases in which the child supplies a 2 p.s./3 p.s. inflection instead of the required 1 p.s., the adult's utterances immediately preceding the child's utterance containing the error included a verb inflected for 2 p.s.:

(21) M: **Quer** isso também?
"Do you want this too?"
Quer um pedacinho?
"Do you want a piece of this?
F: **Quer.**
"You want" (F., 3;02.26)

(from Rubino & Pine, 1998: 450)

Contrary to Borer & Rohrbacher's proposal it is in fact the case that children do commit agreement errors. While this finding is problematic for an approach like the SCH that postulates the *a priori* existence of FCs in ECG, it is compatible with a lexicalist account which views morphological learning as an integral part of the emergence of FCs. Errors have a place in the child's gradual piecemeal learning of new lexical items. FCs' emergence and their consolidation is therefore directly proportional to children's increasingly accurate and larger vocabularies.

In the next section we will consider another account of children's optionality in the expression of verbal morphology which rests on a maturational account of the development of the interface between grammar and pragmatics.

4.3.3. A pragmatic account of OIs

Yet another analysis of the puzzling phenomenon of the OI stage which accounts for the non-target status of ECG in terms of a pragmatic deficit , is the one proposed by Wexler in a series of slightly different versions (Wexler, 1994; 1996; 1998; in press). Here the focus will be on the latest version developed by Wexler together with Schütze (Schütze & Wexler, 1996; Wexler, 1998): the ATOM (Agreement/Tense Omission Model). As the name suggests the core argument of the proposal is that the OI phenomenon is the result of the omission of Inflectional functional projections of Agreement and/or Tense. Wexler (1994) originally proposed that OIs were the result of a missing T category, this however left unexplained some facts about the case-marking of subjects of OIs. If it is only T that can be omitted, and AGR is projected, how can non-nominative Case subjects be accounted for? By contrast, the ATOM makes exactly the right kind of predictions with respect to the Case marking of OI subjects. If either AGR or T can be omitted, OIs will have two different structures:

(21)a. [+AGR, -T] e.g. He like ice cream.
b. [-AGR, +T] e.g. Him like ice cream.

Because in both cases one of the Inflectional FCs is missing, the verb will be non-finite. In the case where AGR is projected , and T is not, the subject will have nominative Case, as expected. If, instead, AGR is missing, but T is present, then the

subject will get default Case assigned to it, which happens to be accusative Case in English.[24]

The occurrence of non-nominative Case subjects does however raise a problem in that it is standardly assumed that subject raising out of spec-VP is motivated by the need for the subject DP to check off its -Interpretable nominative Case feature. Why do subjects like *Him* in (21)b raise at all, if they have no nominative Case feature to check? Wexler (1998) capitalises on Chomsky's (1995) version of the Extended Projection Principle:

(22) EPP is the requirement that a D feature be checked.

This would account for the cyclic movement of subjects of raising verbs where no Case features are checked by the intermediate specifier position of the embedded non-finite I:

(23) Mary$_i$ I [+finite] seems [t$_i$ to I [-finite] [VP t$_i$ like it here]]

What really drives subject raising is therefore not a -Interpretable Case feature on the DP, but rather a -Interpretable D feature on AGR and T that must be checked off and erased by the raising of a subject DP in spec position. Unlike the AGR and T D feature which is -Interpretable, as all target features are, the D feature on the DP is actually Interpretable, this accounts for the fact that it can raise to more than one spec position. If it was -Interpretable, after checking with the lower FC it would be frozen in place and thus would make the DP unavailable to subsequent raising.

This revised interpretation of the EPP has important implications for the explanation of the crosslinguistic variation observed in the occurrence of OIs. Summing up the results of a number of acquisition studies in typologically different languages Wexler (1998) proposes the Null Subject/Optional Infinitive Generalization:

(24) The NS/OI Generalization

> Children in a language go through an OI stage if and only if the language is *not* an INFL-licensed null-subject language.

[24] Default Case is subject to crosslinguistic variation, and it is not always some oblique Case. In German and Dutch, for instance, default Case is assumed to be nominative.

This previously unexplained difference in the existence of OIs between INFL-licensed null subjects and languages where, if null subjects exist at all are not licensed by INFL, can now be related to the parametric variation of the existence of a D feature in AGR. In an INFL-licensed null subject language like Italian, AGR is pronominal (cf. Rizzi, 1994a; 1994b), it *is* D and therefore does not need D, and does not need to be in a spec-head relationship with a DP carrying a D feature. An other way to phrase Wexler's suggestion would be to say that the D feature on Italian AGR is not -Interpretable, as in English, but Interpretable, and therefore any raising to it of a DP carrying an Interpretable D feature is blocked to avoid what Chomsky (1995: 283) calls "'locally superfluous' movement operations".

In addition to assuming that children know about the EPP, as part of UG, and that the EPP is driven by the requirement to check a D feature which is parametrically specified as either -Interpretable (non-null subject languages) or Interpretable (null subject languages), Wexler makes another crucial stipulation: the grammar of children at the OI stage is subject to the Unique Checking Constraint:

(25) The Unique Checking Constraint (UCC)
The D feature of DP can only check against one functional category.

This explains why OIs are only found in non-null subject languages. In such languages both T and AGR need to check a -Interpretable D feature, hence subject DP raising is required to both spec positions. However, because the child grammar must obey the UCC, in order for the derivation to converge one of the FCs must be omitted otherwise there would be a -Interpretable feature which has not been checked:

(26) [AGRP AGR (D) [TP DPi T [VP ti V]]]

In (26) the AGR category has a -Interpretable feature D which has not been checked and erased by the raising of the subject DP to its spec position, such a derivation is not allowed as it would crash. The child is sensitive to the requirement that all -Interpretable feature be checked and erased so that they will be invisible at LF processes, and her grammar cannot allow a derivation like (26). By missing out either AGR or T, an OI results. By contrast, the UCC does not pose any problems to children acquiring a null-subject language, AGR contains an Interpretable D feature which does

not need to be checked by a subejct DP, the D feature is only checked once in spec-TP, the derivation converges and the UCC is satisfied at the same time.

Although Wexler's ATOM explains many puzzling peculiarities of the OI phenomenon, neverthelss it also raises a number of problems, the most significant of which is the optionality in the application of the UCC. The UCC is presented as a principle of child grammar and as such it should not have an optional status, but rather an obligatory one. Because RIs coexist with grammatical finite forms, the UCC does not apply across the board but only optionally. Secondly, Wexler assumes that in a null-subject language AGR and T are always part of the grammar and that the corresponding omission in children acquiring non-null subject languages is not because these FCs are not available in principle, but simply because children's grammars must satisfy the UCC. Although it is true that children learning a language like Italian produce an insignificant number of RIs, nevertheless this approach raises a non-trivial problem. If no Past vs. -Past morpholexical contrasts are operative in the child's grammar, to what extent is it licit to assume that a T category is actually relevant in the child's grammar at this stage? Hyams (1992) comments that concluding from the absence of particular lexical items that a grammatical category is missing from the child's grammar is a "dubious assumption". To which Platzack (1992: 77) aptly retorts that "it seems ... just as dubious to claim there are functional categories present in the absence of any empirical evidence". A more cautious approach to the postulation of FCs is the one adopted by proponents of the Weak Continuity Hypothesis: in their view the assumption that a given FC is present in the child's grammar is warranted only by unambiguous evidence that the child's grammar behaves exactly like adult grammar. Up to that point, the structural description of the child's utterances must be limited to those and only those FCs that are necessary and sufficient to define the child's grammar at that particular point in time.

4.3.4 Structural economy and the Weak Continuity Hypothesis (WCH)

Like the stronger version of the CH, the WCH argues for the availability of UG principles from the onset of acquisition. By contrast, initial availability of FCs is not a precondition of language acquisition in the WCH, but rather the result of the interaction of UG principles and the lexical learning of language-specific morphosyntactic features.

Two major assumptions underlie the hypothesis put forward by Clahsen and colleagues (Clahsen, 1990, 1992, 1996; Clahsen & Penke, 1992; Clahsen, Penke & Parodi, 1993/94; Clahsen, Eisenbeiss & Vainikka, 1994):

(27)a. Head-driven projections:
There is no fixed set of labels for functional projections CP, IP, DP, ... provided by UG; rather *functional projections are feature bundles*, and their properties are determined by their head features.

b. Structural economy:
At any point in a derivation, *a structural description for a natural language string employs as few nodes as grammatical principles and lexical selection require* (Safir, 1993: 12).

(Clahsen, 1996: 131, my italics)

The idea that FCs are feature bundles undermines the notion of FCs as monolithic entities that exist as an all-or-nothing phenomenon. In other words, the acquisition process need not be seen as the acquisition of a given FC AGR, T, or C, but in terms of which, if any, of the features that define the functional head are relevant in the child's grammar yet. Let us take, for instance, the AGR category; the bundle of features that typically define AGR are Person and Number features. If, at a given time in the acquisition process, the Number feature is part of the child's grammar while the Person feature is not, what is the status of the AGR category? Proponents of the SCH, such as Hoekstra & Hyams (1998), claim that what is represented in the child's grammar is an underspecified AGR node, where only Number, but not Person is specified. In the light of (27)a, researchers working in a WCH framework would propose that the AGR node, in the way it is understood in the adult grammar as a bundle of Person and Number features, is not instantiated in the child's grammar yet. What is projected is a functional category F which is only specified for Number.

In the end the difference between the SCH position and the WCH position may reduce to a terminological issue. Considering the recent proliferation of FCs where each feature can ultimately be thought of as projecting into a FC of its own, underspecified FCs or emerging F categories will be virtually undistinguishable. What can still distinguish a SC approach from a WC approach is the theoretical stance on the a priori availability of FCs. For the SCH it is still position-before-feature acquisition: for instance, a T category is by necessity part of the child's grammar a priori, while its feature specification is subject to variation during development. By contrast, for the WCH it is the acquisition of a morpholexical correlate of the FC in question that justifies its presence. This parsimony of structure naturally leads to the issue raised in

(27)b. One of the postulates of the WCH is that it is not desirable to assume more structure than is necessary to account for the data at hand. In other words, if a child's grammar is consistent both with a CP analysis and an IP, or an FP analysis, where F is simply some functional projection marking finiteness, then the analysis which requires less structure is to be preferred over the one that presupposes more functional nodes. Clahsen, Penke & Parodi (1994: 405) distinguish between an analysis that is *compatible* with the data and one that is *supported* by the data:

> "... an analysis is considered to be *compatible* with empirical evidence if the analysis accounts for all the data but assumes phrase-structure positions and/or features that are always empty in the sentences the children produce. ... an analysis is *supported* by empirical evidence if the analysis accounts for all the data and does not assume phrase-sructure positions that are always empty in the sentences the children produce"

The distinction between an analysis that is compatible with the data, and one that is supported by the data is crucial in deciding how much knowledge of FCs a child can be credited with. The position adopted by Clahsen and colleagues is that there is reason to posit a phrase-structure position X only if X is either phonetically realized, or involved in some syntactic process (e.g. as a landing site for a moved element).

A third major tenet of the WCH is that lexical learning is the driving force behind the emergence of FCs. Because UG does not specify that the whole inventory of FCs is available in all languages, the child will have to decide on the basis of the positive evidence available in the input which FCs are instantiated in her language and whether the associated features are strong or weak. For German child language, Clahsen and colleagues have gathered a substantial body of evidence indicating that children's initial clauses lack both AGR and C projections, and that an emerging functional category F simply marking finiteness is available in the earliest phases. Both Clahsen & Penke's (1992) investigation of the Simone corpus (Miller, 1973), and the two studies of seven and five monolingual German-speaking children respectively in Clahsen, Penke & Parodi (1994) and in Clahsen (1996), focus on the crucial role played by the acquisition of the AGR paradigm in the creation of the corresponding functional projection. Clahsen & Penke propose an early stage in the development of Simone's German where finite verbal forms, such as lexical verbs inflected for *-t,* modals and forms of *sein* ("to be"), are typically found in second position, as required by the V2 rule of adult German. Non-finite forms ending in *-en* are in sentence-final position, again confirming to the adult target. However while V2 movement in adult Germain is considered to be movement to C, Clahsen & Penke do not believe that the position to which Simone's early finite verbs move to can count as C. They argue that it

is simply an F category specified for finiteness which is not yet specified for syntactic category. Non-finite forms, which do not carry the [+finite] feature, presumably remain the base-generated V position. It is not until the 2nd person singular marker *-st* is acquired that the child can be said to have acquired the AGR paradigm and the related Person and Number features. Acquisition of the paradigm triggers the realisation of the AGR head where verbs can move to check and erase their strong Person and Number - Interpretable features.

Similar conclusions are reached by Clahsen, Penke & Parodi (1994) and Clahsen (1996) on a larger sample of German-speaking children. Once the Agreement paradigm has been acquired, a corresponding AGRP is created where a specifier position is now available for subjects raised out of the VP. Syntactic evidence for the presence of a newly created AGR projection after the acquisition of the Agreeement paradigm is provided by utterances containing a negative element. The assumption in adult German is that topicalized elements are in spec-CP, the finite verb is in C, and the non-topicalized subject is in spec-AGRS followed by the negative marker situated to the left of VP. In adult German a negative utterance containing a non-topicalized subject will have the word order V-Subj-Neg. At a stage where no AGR projection is available to the child this word order should be unattested, the expected word order being V-Neg-Subj, where the subject has remained in spec-VP to the left of the negative element. This prediction is confirmed by the data in Clahsen (1996) where the adult word order V-Subj-Neg is attested only after there is evidence of productive use of the Agreement paradigm.

The major appeal of the WCH is thus the principle of structural economy combined with an emphasis on the nature of the morphosyntactic input available as positive evidence to the child. In addition to the WCH, a number of other lexicalist proposals have also recently appeared in the literature, notably Tomasello's Verb Island Hypothesis (1992) and Lieven, Pine and colleagues' limited-scope account (Lieven, Pine & Baldwin, 1997; Pine & Martindale, 1996; Rubino & Pine, 1998; Pine, Lieven & Rowland, 1998). The next section illustrates the major tenets of these two approaches to language acquisition.

4.4. Lexicalist-constructivist accounts of early verbs

In sharp contrast with accounts of language acquisition that postulate the existence of syntactic categories from the earliest stages of the child's production,

Tomasello (1992) proposes a radical constructivist alternative in the Verb Island Hypothesis. In Tomasello's view, when children start to produce multiword utterances they do so without any knowledge of either syntactic categories of verb, noun, and determiner, or of semantic roles such as agent or patient. Their initial knowledge about language is highly lexically specific and at most includes categories such as "hitter" and "thing hit" which arise from the child's personal experience with individual predicates. Children's early verbs are islands of information which are not yet systematically organised according to paradigmatic relations. Early knowledge about verbs is best characterised in terms of item-specific information whereby what is known about an individual predicate is not automatically transferred to other predicates, as would be expected if the child could relate individual predicates to an abstract notion of verb category.

Using experimental elicitation techniques with a group of English-speaking two-year-olds, Olguin & Tomasello (1993) showed how at an early stage of language production children's knowledge about SVO word order is associated with a limited number of known verbs and children have difficulties in generalising this information to verbs they have never heard before. At the same time Tomasello & Olguin (1993) showed that the same children had great facility in using nonce nouns into a familiar verb structure. The results of these two studies point to a developmental asynchrony between the emergence of a noun category and a verb category, where paradigmatic relations in the verbal domain take longer to become established.

The principal assumption at the core of the Verb Island Hypothesis is that children's acquisition proceeds via lexically based distributional learning. To begin with children pay attention to individual words, and what they learn about them is word-specific. Because there is no stipulation of the a priori availaibility of syntactic categories, whatever information is gained by the child does not immediately feed into a higher order abstract category. Categories, and in particular the verb category, gradually emerge out of the networking of previously independent, self-contained verb islands. It is only when children begin to notice differences and similarities and start to use forms contrastively that they begin to relate them to a paradigm.

In a recent study on the use of verb morphology, auxiliary verb structures, pronoun case marking, and SVO word order of 12 English-speaking children in the early stages of multiword utterances, Pine, Lieven & Rowland (1998) take Tomasello's argument one step forward. They propose that verb island effects are the by-product

of the interaction between children's distributional learning mechanism and high-frequency markers and lexically and morphologically defined slots. In this sense verb island effects are simply the result of more general lexical learning effects whereby high frequency lexical items in predictable pattern positions are learnt by the child relatively early. Lieven, Pine & colleagues argue that an account of Early Child Grammar does not need to incorporate notions of pre-existing syntactic categories, rather they propose that such categories are the result of a gradual, piecemeal process of lexical learning where children pay attention to statistical regularities in the input both at the morphological level and at the distributional level. Morphosyntactic acquisition is thus the result of item-specific word learning which builds gradually over time. Presumably when the lexical basis reaches sufficiently large dimensions, previously unrelated pieces of information come together in a more systematic fashion. For example, overlap of morphological markers between different verb types will lead to the realisation that there is a category of lexical items that all take a certain morphological shape (e.g. aspectual ending *-ing*, past *-ed*, 3 p.s. *-s*).

4.5. An integrated approach

Lieven, Pine and colleagues do not commit themselves to what exactly is the representation of syntactic categories once they become productive in the child's grammar. What they are more interested in is the a priori requirements that are needed to account for children's early multiword utterances. The position adopted in this work sees syntactic categories as part of the UG inventory that is in principle available to any child from the beginning of acquisition. At the same time we embrace a Weak Continuity approach whereby, although available *in principle*, FCs are not necessarily realised in child grammar from the start. This is where we believe that constructivist positions such as Tomasello's and Lieven, Pine & colleagues' make a major contribution in the understanding of *how* the emergence of syntactic categories is shaped. The notion that a psychologically real mechanism of distributional learning enables the child to pay attention to statistical frequency effects in the input she is exposed to is one that we subscribe to.

A number of computational implementation of what kind of distributional mechanisms may be available to language learners have recently appeared in the literature (Finch & Chater, 1992; Brent, 1993; Cartwright & Brent, 1997; Redington & Chater, 1998). The claim being made here and in other studies is neither that distributional learning is *the only* learning device available to children, nor that it is

applicable to *all* aspects of language learning. The point is rather that a distributional perspective can account for some aspects of early language acquisition, especially those that require the child to deal with the abstract notion of classes and categories. The analysis of C.'s data in chapters 6, 7, 8, and 9 will corroborate empirically some of the claims made here about the relevance of lexical specificity and distributional learning in early child language.

The next chapter introduces the syntax, semantics and pragmatics of Determiner Phrases (DPs) with specific reference to definite and indefinite articles. An overview of developmental studies on the emergence of referentiality and the acquisition of Number and Gender features is also presented, together with a number of predictions for the acquisitional sequence of DPs.

Chapter 5

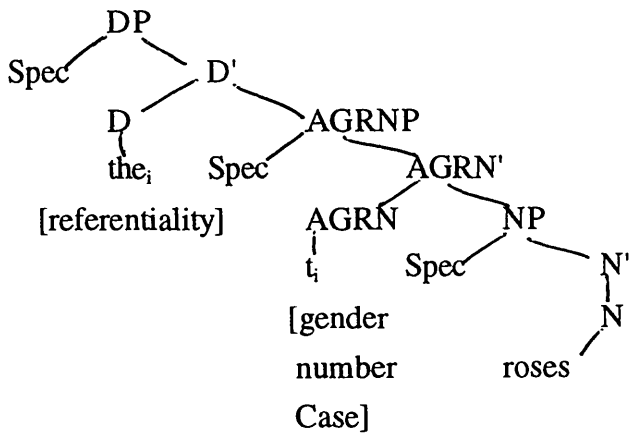
Referential DPs: syntax, semantics and pragmatics

5.1. Introduction

In chapter 4 we outlined some of the basic working assumptions of the MP for formal features and FCs of the verbal domain: AGR and T. In this chapter we will present an analysis of nominals along the same lines. As will be shown below, nominals have a complex syntactic representation that involves a layered structure headed by a Determiner (D). Two distinct sets of formal features are realised by Ds: the Interpretable feature referentiality, on the one hand, and on the other, -Interpretable features such as gender, number, and Case.[1]Two distinct FCs are assumed to be necessary for the checking of nominal features: D, the highest functional head checks referentiality, while AGRN, an intermediate functional projection between DP and NP, checks gender and number Agreement features and Case.

(1)a The roses.

(1)b



The following section will explore issues regarding the nature of definite and indefinite DPs, including singular definite reference, specific and non-specific

[1]Note that Chomsky (1995) treats phi features on nominals such as gender and number as interpretable. Here we adopt Tsimpli & Stavrakaki's (1999) proposal to treat phi-features on nominals as non-interpretable for reasons that will be fully explained in section 5.6.

indefinite reference, generic reference and non-referential DPs, and the pragmatics of definite and indefinite reference.

5.2. Denotation and reference

The issue of reference is central to the syntax and semantics of DPs and it concerns "the relationship which holds between an expression and what that expression stands for on particular occasions of its utterance" (Lyons, 1977: 176). It is important to distinguish here between two terms that are sometimes used interchangeably, but that in fact designate two quite different concepts: *denotation* and *reference*. The crucial distinction between denotation and reference is that the denotation of an expression is utterance-independent, it is part of the meaning which the expression has in the language, and as such it is invariant. For instance the noun *cat* always denotes the same class of animals, regardless of the utterance it is in. Common nouns then denote natural kinds, i.e. classes whose members have the same nature or essence, while proper names denote individual entities. As such, common nouns do not have reference, but they may be part of a referring expression.

While denotation is utterance-independent, reference, by contrast, is utterance-dependent, in that what is referred to by an expression strictly depends on the context of the utterance. While the noun *cat* denotes a natural kind of four-legged furry animals with a tail, the DPs *the cat*, in an utterance like (2) below, clearly refers to a certain individual cat:

(2) The cat has not been in all day.

What makes the expression *The cat* in (2) a referential expression is the use of a determiner, in this case the definite article *the*. It is in fact a characteristic of determiners, such as the definite article, that their primary function is that of determining the reference of the nouns with which they occur.

5.3. Referentiality and D chains

Longobardi (1994: 633) defines deteminers as operators binding a variable whose range is the extension of the natural kind denoted by the head noun. He proposes that the reference of a given nominal expression results from the checking of the referentiality feature in D. Longobardi also assumes that the referentiality feature

may have a [+referential] value (+R), and a [-referential] value (-R). When the value is +R, the referentiality feature is universally checked if D is in a chain containing an object-referring expression (i.e. a noun or a proper name). Conversely, when the value is -R, the referentiality feature is universally checked if D is in a chain *not* containing an object-referring expression (i.e. a common noun). Moreover, Longobardi proposes that the strength of the ±R feature is subject to crosslinguistic parametrisation, and it is strong in Romance languages, thus forcing movement from N to D before Spell Out wherever possible, and weak in Germanic languages where movement from N to D is procrastinated until LF.

The D position may be generated as containing a pronoun, a lexical determiner, or it may be empty. When D is filled by a pronoun, +R is checked off by a single-membered, object-referring chain. If D is occupied by a lexical determiner, the +R value can be checked off only by establishing a chain with an object-referring expression (a proper name). The possibility to form a chain between a lexical determiner and a proper name however, is only available in those languages where the definite article is expletive, such as Italian, but not in those languages where the definite article cannot perform an expletive function:[2]

(3) La Maria è venuta a trovarci ieri sera.

(4) *The Mary came to see us last night.

When D is base-generated as empty, raising of a proper name from N to D is required to check the +R feature. Such movement takes place before Spell Out in Italian where the ±R feature is strong, but it is delayed until LF in English where the ±R feature is weak.

If the value of the referentiality feature is -R, checking is only possible if a chain is formed between the D position and the N position containing a kind-denoting expression such as a common noun. The only two situations in which -R can be ckecked are when D is occupied by a lexical determiner, and when D is empty. If it is occupied by a personal pronoun (an object-referring expression), no checking is possible as stated above. If D is occupied by a expletive article , and N is occupied by a kind-denoting expression in N, i.e. a common noun, -R is checked by the formation of a chain between D and N, and the generic reading results.

---

[2]Note that the availability of a definite article before a proper name in (3) is subject to dialectal variation, and is commonly found only is some northern varieties of Italian.

(5) Il castoro è un roditore./I castori sono roditori.
"The beaver is a rodent./Beavers are rodents"

If the lexical determiner is a non-expletive article, it checks -R in a single-membered chain, however because a determiner is an operator, it still needs to bind a variable, in this case the extension of the natural kind denoted by a common noun in N (definite reading).

(6) Il castoro che ho visto al fiume era molto grosso.
"The beaver that I saw at the river was very big."

If D is base-generated as empty and -R is strong, as in Italian, the D position will be interpreted as existential and will be subject to the a lexical government requirement. The existentially-interpreted null D operator will also need a variable with a range to bind, i.e. a commoun noun:

(7) Quest' anno non abbiamo visto castori al fiume.
"This year we have not seen beavers at the river."

In English an empty D position can also give rise to a generic interpretation if, at LF, the kind-referring expression is raised from N to D:

(8) Beavers are rodents.

In sum, the D position is the position reserved to object-referring expressions, such as personal pronouns and proper names (in Italian), and determiners whose function as operators is to bind the range of the natural kind denoted by common nouns. In the next section we see how the distribution of definite and indefinite articles is also semantically and pragmatically constrained by specificity and familiarity.

5.4. DPs: reference, specificity and familiarity

Depending on the determiner selected in the D position, there will be DPs with different referential interpretations: specific definite reference (*the cat, the man standing over the door, the sun,* etc.), specific and non-specific indefinite reference, and generic reference. Specific definite reference is typically associated with proper names, personal pronouns (e.g. *I, you, she, he, we,* etc.), and definite DPs (*the cat, the boys, the moon*, etc.). While proper names and personal pronouns are intrinsically

referential, definite DPs are referential by virtue of the definite article which identifies a specific unique entity, or entities, in the utterance context. Indefinite DPs, on the other hand, may be specific or non-specific, depending whether it is the speaker's intention to identify a particular member of a class, or any member of a class.

At the core of the notion of specificity is the abstract system of classes and class membership, as well as the idea of any class member (Maratsos, 1976). In particular, specific reference presupposes that the speaker be familiar with the notion of unique member of a class. A unique member of a class is such by virtue of some distinctive property that sets it apart from all the other members of its class. In situations where the specific referent is physically present this distinctive property may be of a perceptual nature. In other words, the referent is unique because it physically occupies a unique spatio-temporal setting that that and only that referent occupies at a given time. For example, in a request such as *Please give me the book* indicating a book, the speaker is singling out a specific book that is so by virtue of being physically present and being indicated as the one and only book in the speaker's intention. However a specific reference can also be made even if the referent in question is not actually phisically present. By saying *I saw a girl being mugged yesterday* the speaker is making specific reference to a girl whose unique property is that of having been seen by the speaker while she was being mugged.

By contrast, non-specific reference is associated with the notion of any member of a class. The non-specific referent may be a real one as in *Give me a fork, please* where the speaker is not asking for any particular fork but for any metallic object as long as it belongs to the class of forks. Alternatively, the referent may be of a more prospective or hypothetical nature as in *I want to buy a new house* or *I wish I had a car*, where no specific house or car is intended but any house or car as long as they are members of the respective house and car classes.

The other crucial notion involved in the selection of the appropriate article for expressing specific vs. non-specific reference, is the discourse property of presupposedness. In other words, the speaker must take into account the listener's knowledge when identifying a referent. If a referent is [+specific] and is not known to the listener, the speaker will have to use an indefinite article on first mention. If, conversely, the specific referent that the speaker has in mind is also known to the listener, then the speaker can use a definite article.

Researchers have proposed different interpretations of the initial availability of the notion of specificity to young children. Piaget (1962) claimed that initially children have difficulties with the notion of class and class membership, and as a consequence they are incapable of dealing adequately with the distinction between specific and non-specific reference. Similarly Bruner et al. (1966) suggested that in the early stages children rely predominantly on iconic representations and do not have the mature competence to abstract from imagic representations to form the notion of abstract class and of any member of a class. However, in a series of experimental studies with English-speaking children between 2;8 and 4;11, Maratsos (1976) provided substantial evidence to show that, contrary to what had previously been thought, even children younger than 3 do possess a very sophisticated competence with specific vs. non-specific reference.

Starting from the discovery of striking similarities between diverse creole languages Bickerton (1981, 1984) argues that there must be a universal species-specific biological program for language acquisition if we are to account for these otherwise unexplained similarities. He claims that one of the innate distinctions available to children is the specific/non specific distinction. Cziko (1982) reviews a number of studies (Brown, 1973; Maratsos, 1976; Warden, 1976; Karmiloff-Smith, 1979; Garton, 1983) in the attempt to evaluate Bickerton's (1981) claim that the specific/non-specific distinction must be biologically programmed. Although the original aim of most of the studies examined was to investigate children's sensitivity to familiarity rather than to the specific/non-specific distinction, Cziko concludes that there is good direct or indirect evidence from most of these studies that even very young children are indeed able to make the appropriate distinction.

Assuming that it is indeed the case that children are sensitive to the specific/non-specific distinction from very early on, what about sensitivity to familiarity? Previous research has shown that the ability to take into account the listener's point of view develops over time and it is not always part of the child's initial communicative competence. When choosing an appropriate article the speaker must minimally obey two semantic-pragmatic constraints: firstly, as mentioned above, the speaker must decide whether the referent she wants to talk about is specific or non-specific, i.e. if it is of a particular, unique member of a class that she wants to talk about or of any member of a class. Secondly, the speaker must decide whether the listener shares the same knowledge about the referent in question; in other words, the speaker must know whether she can presuppose that the listener is familiar with the referent in

question or not. The combination of specificity/non-specificity (where the relevant feature is ± S) and familiarity/unfamiliarity (where the relevant feature is ± F) gives rise to the four possible combinations: [+S;+F], where the referent is specific for the speaker and is familiar to the listener; [+S;-F] where the referent is still specific for the speaker but not familiar for the listener; [-S;-F] where the referent is non-specific for both speaker and listener, and finally [-S;+F] where the speaker is non-specific for the speaker but is familiar to the listener. Table I below exemplifies the four conditions and the type of article required whether definite, indefinite or null. The distribution of articles applies to both English and Italian although the null option is only available in English.[3]

contexts	definite article	indefinite article	null article
[-S;-F]	-	x	x
[+S;-F]	-	x	x
[+S;+F]	x	-	-
[-S;+F]	x	x	x

Table I. The distribution of aticles in the different semantic-pragmatic contexts

As explained by Brown (1973; 345-347) there are a number of contexts where the speaker can use a definite article for a referent in the certainty that the listener will also be familiar with it: when the referent is a) unique for all (e.g. *the sun, the sky*), b) unique in a given setting (e.g. *the couch, the floor*), c) uniquely salient for a given social group (*the teacher, the queen*), d) made salient by pointing, nodding, spotlighting, e) made salient by stimulus characteristics that capture attention (*the dog, the explosion, the motor*), f) specific by entailment (*the hand, the engine, the husband*); g) specific by definition (*the last sentence, the best)*; h) specified by prior utterance (*the stupid cat*).

It is clear from this list that in seven out of eight contexts in which the definite article is used, the listener's familiarity with the referent can be presupposed on extralinguistic grounds. The definite article is used exophorically rather than endophorically such as in the last context, where its use is justified by the fact that a prior utterance has introduced it as a shared discourse referent, therefore reference to it can now be made anaphorically. In all the other cases encyclopedic knowledge or the

[3] Recall that in Italian the distribution of null articles is limited to a small number of lexically-governed positions and the only interpretation available for DPs with a null determiner is the existential reading. The specific definite reading, as well as the generic reading always requires the presence of an overt article

situational context provide sufficient evidence for the referent to be treated as unique and familiar both to the speaker and the listener.

An exception to the use of the definite article when the familiarity of the referent is inferred by the situational context is that of naming. Brown (1973) observes that when pointing and naming something new, parents and children often say things like *That's a bear, this is a crocodile.* However, if the referent is phisically present and clearly available to both speaker and listener, the use of the definite article would be expected, not the indefinite. Brown accounts for this by arguing that, although the object is physically present and thus specific for both participants in the conversation, nonetheless it does not yet exist *by name* for the listener; "In this sense they do not exist as what Kartunnen calls 'discourse referents' though the objects themselves, without names, are specific for the listener" (Brown, 1973: 348). Maratsos (1976) points out that, although this may well be true, there are cases in which the referent is clearly available to both speaker and listener as a discourse referent and yet the indefinite article is still used: "'It is, after all, only a bear' or 'The fact that it is a bear should not affect us'" (Maratsos, 1976:7). In these cases Maratsos is more inclined to treat these DPs as instances of a non-specific referent where the speaker, although pointing to a specific bear, is only concerned with it as a non-particular member of the class of bears. Thus while naming for Brown should be considered as [+S;-F], for Maratsos it would be considered as [-S;+F], although he also entertains the possibility that naming may not be accurately described by any of the four categories in Table I at all but may need a classification of its own.

A clear-cut case in which the indefinite article is required is when the referent is specific for the speaker but cannot be specific for the listener as he is not yet familiar with it [+S;-F]. Familiarity cannot be inferred by any of the contexts from a) through h) above. For instance, if a speaker says *When I arrived home yesterday a woman was waiting for me by the door,* she is deliberately using an indefinite article to talk about the specific referent *woman* as she knows that, although it is specific and familiar to herself this referent is not familiar to her listener in that he may very well know nothing about what happened to her the day before. In English in such a [+S;-F] context the null article is also available for mass nouns and plurals: *I was eating pasta/tomatoes in the kitchen when the phone rang.*

When a speaker does not intend any specific member of a class but any member of a class the article required is again an indefinite or, in English, its null

counterpart for mass nouns or plural nouns: *Give me a glass, I need shoes for the party tonight, I want money for new clothes.*

Following Thomas (1989) I have included generics in the [-S;+F] category. Both the definite and the indefinite article can be used in English and Italian; in addition for English the null determiner option is available with mass nouns and plurals: *A dog is always better than a cat, The Italian wolf will soon become extinct, Crocodiles are dangerous animals, Fruit is an essential part of the diet.*

5.5. Previous studies on the mastery of specificty and familiarity

At this point it is obvious that the notion of specificity and that of familiarity are strictly interdependent and that it is impossible to master knowledge of the article system without having acquired communicative competence in both. Since there seems to be good evidence that children are indeed sensitive to the specific/non-specific distinction, experimental studies on the acquisition of articles have largely focused on the availability of the notion of familiarity, in particular on the ability of young children to supply an indefinite article for the [+S;-F] context. In principle two types of errors can be expected if children have not yet fully mastered the notion of familiarity: they could use an indefinite article although the referent is familiar to the listener either because introduced by a previous utterances or because of the situational context. In this case the indefinite would be erroneously employed in a [+S;+F] context where only the definite would be correct. This type of error has been defined in the literature as an 'incoherence error' (Emslie & Stevenson, 1981). Although in experimental conditions these errors are unreported, in spontaneous production Brown (1973) reports 14 incoherence errors in the combined article output of the three children in his corpus.

The second type of error, the 'egocentric error', can occur when the child mistakenly assumes that the listener shares his point of view and is thus familiar with the referent just like he is, this is the case in which a child employs a definite article for a [+S;-F] context where an indefinite is required instead. Several studies have reported this kind of 'egocentric errors' (Warden, 1976; Maratsos, 1976; Emslie & Stevenson, 1981; Power & Dal Martello, 1986; Thomas, 1989). It is as if children were only capable of taking into account their own intentions when speaking and thus use a 'speaker-centred' rule rather than a 'listener-centred' rule (Power & Dal Martello, 1986). Brown found 14 instances of egocentric errors in his naturalistic data and in

experimental conditions Maratsos (1976) reports that only four three-year-olds, out of twenty, and eleven four-year-olds, out of twenty, made no more than one error in the [+S;-F] context by supplying a definite article instead of an indefinite one, and at the same time produced definite articles where required at least 75% of the time. Similar results were obtained by Warden despite the fact that, unlike in Maratsos' experiments, the children were not talking to the experimenter, whom they could have assumed to be familiar with the materials used, but to a third person who could not see the materials the child was looking at, and was therefore genuinely unfamiliar with them.

Emslie & Stevenson (1981) obtained a significant improvement in the children's performance by making some adjustments to the experimental conditions. Their suspicion was that the high number of incorrect definite articles might have been biased by the fact that the stories were too implicit and by the children not actually talking to the other subject but to the experimenter. The stories were simplified and it was made very clear to the children that they had to talk to the other subject who could not see the cartoon story and not to the experimenter. Table II summarizes Emslie & Stevenson's results:

	1st mention	1st mention	2nd mention	2nd mention
age group	Indefinite	Definite	Indefinite	Definite
3 years	84	13[4]	4	96
4 years	85	15	0	100

Table II. Percentages of definite and indefinite reference on first and second mentions in Emslie & Stevenson (1981)

The improvement is clear, but may still not truly reflect whether the children are in fact really using an adult-like, listener-centred strategy. As Power & Dal Martello (1986) showed for a group of fifty monolingual Italian children aged between three to five, it may well be the case that the children can perform satisfactorily and still not employ a listener-centred but a 'discourse-centred' strategy. In order to disambiguate the two possibilities they repeated Emslie & Stevenson's experiment twice, the second time they had the children narrate the stories to two different blindfolded listeners, the results are shown in Table III:

[4] The other 3% of responses were inaudible.

	1st mention	1st mention	2nd mention	2nd mention
narration	Indefinite	Definite	Indefinite	Definite
1st narration	61	39	10	90
2nd narration	40	60	6	94

Table III. Percentage of definite and indefinite reference on first and second mentions for stories told to two listeners in turn in Power & Dal Martello (1986)

As shown in Table III, although the use of the definite article on first mentions is still significantly lower than on second mentions, a noticeable difference is observed between the usage of the definite on first mention between the first and the second narration (an increase in error rate of 21%). This difference is to be expected if some children are adopting a speaker-centred perspective whereby, although the information is new to the second listener as it was to the first listener, it is now known to the speaker who then assumes that the information will be equally familar to the second listener, which is obviously not the case from a listener-centred point of view. Children as old as four still commit a large number of egocentric errors whereby they assume a speaker-centred perspective, a fact that neither Warden nor Emslie & Stevenson could show in their studies. Nonetheless, the significant difference in the use of definite vs. indefinite articles between first and second mentions in both the first and second narration indicates that a significant proportion of the children are sensitive to the listener's point of view and use it to decide whether a definite or an indefinite article should be used.

Thomas (1989) following Cziko (1982) suggests an alternative explanation for the incorrect use of definite articles in [+S;-F] contexts. She observes that learners of English as a second language (ESL), like children learning English as their first language, erroneously extend the use of the definite article from [+S;+F] contexts to [+S;-F] contexts but never to [-S;±F].

Admittedly, data from [-S;+F] contexts is harder to come by, especially in first language acquisition where the production of generics is low given young children's tendency to situate conversations in the 'here and now'. Instances of [-S;-F] contexts are however more common and Thomas claims that errors involving a definite article instead of an indefinite article are unattested either in L1 or in L2. In the case of adult learners, the egocentricity argument obviously loses much of the appeal it has for child language. In the attempt to give a unified account for the L1 and the L2 data, Thomas concludes that the learner's initial hypothesis, whether L1 or L2, is that the English definite article *the* is associated with the [+S] feature and the indefinite *a/an* with the [-

S] feature. If this were the case, while we would not expect to find definite articles in [-S] contexts, similarly we would not expect to find indefinite articles in [+S] contexts. Although Thomas is mainly concerned with showing how the overproduction of *the* depends not so much on egocentricity but on the association with the [+S] feature, she does not comment further on the association of [-S] with *a/an*. Naturalistic data from Brown (1973) provide counterevidence to the supposed initial association of the indefinite with [-S]. Brown's three children do in fact make a number of errors by using an indefinite in a [+S;+F] context, fourteen such errors are reported together with fourteen errors in which the definite is incorrectly used in [+S;-F] contexts. If Thomas and Cziko's claim has any descriptive and predictive power it is only in an asymmetrical fashion; while it may well capture the data for the definite article, errors with the indefinite in [+S] contexts remain unaccounted for.

Although the available experimental evidence on comprehension and elicited production of determiners in children is indicative of a certain trend to overuse the definite article for a referent on first mention in situations in which mutual knowledge cannot be inferred, the limitation of most of these studies lies in the elicitation technique where isolated DPs are often elicited in response to questions and do not form part of a discourse continuum. In the attempt to obtain a more realistic picture of children's communicative competence, recent studies have investigated children's referential expressions in narratives (Bennet-Kastor, 1983; Bamberg, 1987; Wigglesworth, 1990; Kail & Hickmann, 1992; Orsolini, Rossi & Pontecorvo, 1996; Hickmann, Hendriks, Roland & Liang, 1996). Narratives elicited through picture stories such as *Frog, where are you?* (Mayer, 1969) have allowed researchers not only to assess sensitivity to familiarity, but also the roles of topicality and pragmatic predictability in children's use of referential expressions. Bamberg (1987), Kail & Hickmann (1992) and Karmiloff-Smith (1985) reported that for younger children (up to the age of 3-4 for Bamberg, and 6 for Karmiloff-Smith and for Kail & Hickmann) a referent's topicality, i.e. its salience in the narrative plot, plays a fundamental role in the choice of referential expression. Overall children are more likely to use an indefinite form when introducing a secondary character than a main character. Kail & Hickmann (1992) also report that the boy in the frog story has a privileged status, being the protagonist of the story and the only human referent. Children in all their three age groups (6-year-olds, 9-year-olds and 11-year-olds) were more likely to use an definite form rather than an appropriate indefinite form when introducing the boy in situations in which the listener was blindfolded and therefore did not share the child's knowledge of the referent. The decisive developmental shift occurs when children start to rely on pragmatic

predictability, i.e. the pragmatic link between the current utterance and the context representation, rather than on topicality to choose the appropriate referential expression.

In this section we have illustrated how the notion of specificity and that of familiarity are inextricably related to each other in the use of articles. The selection of a definite or indefinite article when a speaker decides whether he or she wants to speak of a specific member of a class, or any member of a class, is inevitably constrained by a pragmatic discourse rule which requires the speaker to take into account the listener's knowledge about the referent in question. The literature on language acquisition has given evidence for children's early mastery of the semantic specific/non-specific distinction. At the same time, there is evidence that children younger than three have not yet fully mastered the pragmatics of article use, and they still make a number of egocentric errors.

The next two sections summarise some of the findings on the acquisition of formal gender and number features in relation to the acquisition of definite and indefinite articles.

5.6. The emergence of gender

As mentioned in section 5.1., in addition to the Interpretable ± Referential feature, DPs also include -Interpretable phi features such as gender, number and Case. In Chomsky (1995) phi features on nominals are treated as interpretable, while phi features on verbs are considered as non-interpretable. In this work we have adopted Tsimpli' and Stavrakaki's (1999) proposal to regard phi features on nominals as -Interpretable. The argument is that agreement features in nominals are realised on all heads of the nominal projection as a result of a N-to-D head Chain formed through percolation, these are resumptive features that are erased at LF as -Interpretable (Manzini, 1995). The heads involved in the Chain, by contrast, will survive at LF depending on their remaining feature specification, thus N and D will still be visible at LF, because of their inherent interpretability, but AGRN will not, since it only includes resumptive -Interpretable features that have been erased by the checking process.

In languages where grammatical gender is morphosyntactically realised, such as Italian, a combination of semantic, morphophonological and syntactic cues can be used by the child to determine the gender of a noun and its accompanying determiner. Some researchers have considered the semantic cue as the primary tool the child uses

in the initial stages of acquisition when her grammar is still semantic-based (Brown, 1973; Gleitman, 1981). Levy (1988) calls this position the "meaning first" view, the basic idea is that in order to establish linguistic gender the child will draw on her knowledge of natural gender according to which animates are divided into males and females. As for inanimate objects, the child could potentially classify them into masculine or feminine gender according to some connection they have with masculine or feminine animate entities. Thus, skirts, lipsticks and blouses will be treated as linguistically feminine because of their association with female human beings; by this rationale trousers, pipes and beards would be treated as masculine. Even if we conceded that this strategy could be successful, there remains a large number of inanimate objects and abstract concepts that cannot be easily associated with either natural gender. Moreover, if this strategy were indeed adopted by the child, we would expect him to perform significantly better in the attribution of linguistic gender to animate nouns where the semantic cue of natural gender would point him in the right direction. In an initial semantic stage the child should perform randomly in the attribution of linguistic gender where there are no clear natural gender cues. For instance, it is difficult to see how a chair, a table, a pen or a curtain could be thought of as more of a masculine or a feminine object.

Levy (1983, 1988) suggests as an alternative to the "meaning first" view the "formal" view. According to this proposal the child would proceed to the gender classification of nouns and determiners via a distributional analysis of formal regularities in the input. Salient, transparent, constant morphophonological cues in the input would be recorded by the child and used accordingly to sort out nouns and determiners belonging to different linguistic gender classes. Paradoxically the child would not even need to be aware of the notion of gender, he would only need to notice that there are two (or more, in languages where there are more than two genders) separate classes with different morphophonological patterns in which nouns and determiners, and modifiers, can be divided.

The fact that knowledge of natural gender need not be a prerequisite for gender classification is corroborated by the cognitive developmental literature on the subject. The development of gender relies on two other major notions: gender identity and gender constancy. The former refers to the understanding that every living being is either masculine or feminine; the latter refers to the fact that being male or female is an intrinsic characterisitc of the living being in question which, with few exceptions, remains unaltered throughout his or her life. Studies in the development of gender

identity and constancy report difficulties in children under the age of 3 and up to the age of 6 (Piaget, 1968; Gouze & Nadelman, 1980). Most researchers claim that gender identity for self is not fully established before 3 years of age, and that gender identity and constancy for others typically develops only after gender identity for self has emerged. In other words, if 2-year-olds have not mastered the notion of gender identity it is unlikely that their approach to linguistic gender classification will be informed by such a semantic principle. Of course it is also true that children before 3 do notice gender differences and that they may be guided by some heuristic principle of this sort, but it does not appear very likely.

The "formal" view of gender acquisition simply requires children to notice morphophonological, and possibly syntactic, regularities in the input, a type of sorting task which is not beyond the cognitive development of a 2-year-old. Levy (1983) reviews acquisitional data from several languages to assess the role played by semantic and formal features in determining the gender of nouns. The languages examined include German, Russian, Polish, French and Hebrew. The author has found no cue that children are initially using a semantic strategy to identify linguistic gender. On the contrary she presents convincing crosslinguistic evidence suggesting that well before the age of 3 children use morphophonological regularities in the input to solve the gender puzzle. There is no significant difference between the children's performance with respect to animates and inanimates which one would expect on the assumption that children make use of the notion of natural gender. Overgeneralizations are made by analogy with similar morphophonological forms, which is neatly accounted for by a formal/distributional analysis of the input.

In an attempt to tease apart grammatical, natural and phonological cues to the acquisition of gender, Karmiloff-Smith (1979) conducted a series of five experiments on a population of 341 monolingual French-speaking children between the ages of 3;2 and 12;5. The focus of the experiment was on what type of cues (syntactic, natural, phonological) children would use to establish the gender of the elicited determiners. The stimuli were 39 nonsense words, obeying the rules of phonemic combination in French, matched with pairs of identical drawings of imaginary objects, imaginary animals and anthropomorphous male and female figures which only differed in colour. Different combinations of cues were provided in the various experimental conditions; in one of the tests the article and the noun suffix were consistent (e.g. *un bicron, une bravaise*), in other two the cues were in conflict (e.g. the natural gender of the referent was in conflict with the suffix on the noun: *bicron,* a noun with a typically masculine

ending, referring to a feminine picture; the gender of the article provided was inconsistent with the gender of the noun: *une bicron, un bravaise*). In one of the remaining two experiments there was no additional cue apart from the suffix on the noun (e.g. *deux bicrons, deux bravaises*); in the other the child was presented with a masculine noun accompanied by the appropriate indefinite article, (e.g. *un bicron*) and was asked to invent a name for the equivalent female picture.

Karmiloff-Smith's results show that children as young as 3 can and will use phonological cues on the noun suffix to determine the gender of the corresponding elicited article. Obviously, when natural gender and phonological cues are consistent it is impossible to say which one of the two cues plays a fundamental role in the attribution of gender. However Karmiloff-Smith's findings show that in the absence of a natural gender cue, the phonological cue alone provides sufficient information for gender attribution. Even more striking, in the absence of a clear gender indication on the noun suffix (e.g. *un coumile*) the indefinite article is used by the children to decide on the gender of the noun and on the corresponding definite article (although to a lesser degree of accuracy; under-5-year-olds privilege the suffix on the noun rather than the accompanying indefinite article to determine the gender of the appropriate definite article).

In those cases in which natural gender and phonological cues are in conflict, Karmiloff-Smith reports that "at all ages in the case of discord, the tendency is to use the phonological cue for providing the gender of the article accompanying a noun and not the natural gender cue" (Karmiloff-Smith, 1979:161). The conclusion seems to be that although children can take natural gender into account in cases of no inconsistency, nevertheless there is a strong tendency to rely almost entirely on phonological procedures, with syntactic information becoming more important as the children get older.

In a similar experiment to Karmiloff-Smith's, Pérez-Pereira (1991) tested knowledge of gender on a sample of 160 Spanish-speaking children ranging from 4 to 11 years of age. His results strongly support a "formal" view of gender acquisition whereby children privilege the intralinguistic syntactic and morphophonological cues over the extralinguistic cue provided by the natural gender of the referents. Three main points emerge from Pérez-Pereira's findings: a) the feature [+animacy] does not seem to play a fundamental role in gender attribution as it was the case for Karmiloff-Smith's subjects; b) the larger the number of concordant cues the easier it is for children to

determine the gender of nouns and therefore choose the appropriate determiner; c) as children grow older syntactic cues become gradually more important than morphophonological cues. When children learn that syntax can override phonology they begin to rely more on syntactic cues. When faced with conflicting cues such as in *el programa,* where the noun suffix is typically feminine but the accompanying article is masculine, Spanish children as young as 4 will rely on the gender of the determiner to establish the gender on the noun.

In Karmiloff-Smith's experiment French-speaking children under age 5 privileged the noun suffix to establish the gender of the noun in question rather than basing themselves on the gender of the determiner. In an example like *une bicron,* where the noun ending is typically masculine but the determiner is feminine, children younger than 5 tended to use the cue on the noun suffix to decide its gender rather than the determiner. As the youngest subjects, 4-year-olds, in Pérez-Pereira's experiment already privilege a syntactic strategy, we have no access at the earliest stages of emergence of the grammatical feature gender, and thus no way of telling whether for Spanish children, as for French children, the priority of syntactic over morphophonological cues is a phenomenon which emerges subsequently to a phonological strategy. Clark (1985) reports observations by Montes Giraldo (1976) on two Spanish children who used **una fola* for *una flor* "a flower" and **la mujala* for *la mujer* "the woman" (2;10), and **una mana* for *una mano* "a hand" (2;11). In overregularizing the noun ending to agree with the determiner the child is clearly relying on the article as gender indicator. Although this is only anecdotical evidence there is some indication that a syntactic strategy may well be in place from the earliest stages of gender acquisition in Spanish-speaking children.

Why should it be that Spanish children "discover" the syntactic function of determiners before French children? Along the lines of what Cipriani, Chilosi, Bottari & Pfanner (1993) have suggested for the earlier acquisition of the feminine singular article *la* over its masculine counterpart *il,* the suggestion here is that the higher one-to-one correspondence between *la/una* and an *-a* suffix on the noun, and between *el/un* and an *-o* suffix on the noun make the morphosyntactic concord between noun and determiner more salient in Spanish than in French[5]. In Spanish, unlike in French, there are a default *-o* suffix for masculine nouns and a default *-a* suffix for feminine nouns,

[5] Cipriani *et al.* (1993) explain their subjects' earlier acquisition of the definite feminine article *la* over its masculine counterpart *il* with a vowel harmony strategy between the determiner and noun suffix. In other words, the transition from a phonetically indifferentiated placeholder to the feminine article would be facilitated by the concord between the article's vowel and the noun suffix.

in addition to a number of exceptions (e.g. *la foto, la moto, la mano; el dia, el planeta, el yoga,* etc.) and a number of other suffixes for masculine and feminine gender[6]. In French there are no default gender suffixes, however we do find a number of one-to-one correspondences between suffixes and gender. For example, on the basis of distributional facts alone the French-speaking child can figure out that suffixes like *-ice, -ère* and *-euse, -elle, -eresse* are associated with feminine nouns, and suffixes such as *-eur, -eau* are typically masculine. Although morphophonological regularitites are good predictors for gender in French, there is no single suffix which can be treated as a default gender marker for either masculine or feminine. Therefore in Spanish, but not in French, the agreement between article and noun is made extremely visible by the binary correspondence between feminine article and *-a* noun ending, and masculine article and *-o* noun ending. The morphophonological cue reinforces the syntactic cue of grammatical agreement thus making the latter available to be incorporated in the child's formal analysis of the input.

Given the different situation in the two languages it comes as no surprise that Spanish-speaking children should have noticed the gender indicating function of the determiner earlier than French children, the transparency, regularity and consistency of the concord between article and noun suffix in Spanish all contribute to make it a strong cue for the attribution of gender and the instantiation of syntactic agreement between the Determiner and the Noun.

For Italian, Cipriani et al. (1993) report an initial stage in the acquisition of the article system characterised by the prevalent omission of determiners, followed by the appearance of a generic morphological placeholder, subsequently by phonetic approximations and ultimately by full forms. Generic morphological placeholders are undifferentiated phonetic segments which occupy the position of free morphemes, in this case the article's, but which have no morphophonological consistency. Phonemes such as [en], [e], [ ], [ò] and [a] have been reported to serve as generic determiners in the speech of the six children analysed by Cipriani et al. They appear indiscriminately with nouns of either gender in both singular and plural forms. In a subsequent stage, generic placeholders are gradually replaced by phonetic approximations, e.g. [i] for *il* and [a] for *la,* until full target forms appear. Cipriani et al. remark upon the significance of the undifferentiated placeholder, which they view as a syntactic precursor of the

[6] The following suffixes are masculine: *-aje, -or, -án, -ambre* or any stressed vowels. E.g. *el paisaje, el rigor* (but *la flor, la labor*), *el desván, el enjambre, el sofá, el café.*
The following suffixes are feminine: *-ión, -dad, -tad, -tud, -umbre, -ie, isis, -itis.*
Nouns ending in *-ista* are invariable: *el/la ciclista, el/la periodista.*

article in its own right, by emphasizing its independence from the morphophonological realization of the article. For the authors the placeholder is the spelling out of the syntactic function of the Determiner.

Although Cipriani et al.'s hypothesis is certainly plausible, there is no a priori need to assume that MPHs show anything more than the child's perceptual sensitivity to phonological and prosodic cues in the linguistic input, whereby nouns, which are more salient than determiners, both semantically and phonologically, are reproduced by the child more consistently. Determiners, on the other hand, are less phonologically and semantically salient, but are also very frequent in the input and their distribution is highly regular and predictable, hence the child's initial attempt to reproduce some segment that is statistically very frequent in front of nouns.

5.7. The emergence of number

As illustrated above, in principle gender assignment does not require the semantic notion of natural gender. In fact on the basis of distributional facts alone a child could classify nouns according to the various genders successfully without being necessarily aware of the notion of feminine, masculine or neuter gender. All is required of the child is to notice the existence of different morphophonological classes of words and their distributional behaviour.

As for number, some semantic notion of singularity and plurality must be in place before singular and plural forms of determiners, nouns and adjectives are used correctly. While both gender and number manifest themselves as a variable feature of determiners and other modifiers (gender and number agreement), gender is also an intrinsic feature of nouns and therefore does not change according to other extralinguisitic or intralinguistic variables (gender attribution); number, on the other hand, varies according to the referent it is associated with.

Müller (1994) distinguishes between *number specification* and *number agreement,* the former refers to the process of determining the reference of nominals, i.e. to decide whether they are singular or plural; the latter concerns the morphosyntactic concord between a nominal and its modifier(s). It is clear that the two, number specification and agreement, do not necessarily go together . Number agreement in the noun phrase could be correct but the number specification of the

referent could be wrong, e.g. *un* (definite article, masculine, singular) *bimbo* (masculine, singular noun), "a child", for a plural referent. In other words, it is possible to treat the two separately.

Number marking is a fundamental morphological unit across natural languages and the widespead use of such marking is suggestive of the primary need to signal a distinction between one vs. more than one, i. e. between singularity and plurality, in any human linguistic system including presumably that of young children. This claim seems to be corroborated by the literature on language acquisition; in a survey of the chronological acquisition of a number of grammatical morphemes in English, Brown (1973) reported that despite the homonimity between the allomorphs of the plural marker */s/*, the present tense third person singular affix, and the possessive, the plural morpheme is acquired much earlier than the other two markers, probably due to its priority in the semantic hierarchy. Number is also supposed to emerge as a grammaticizable notion earlier than gender because of its semantic primacy (Leonard, Bortolini, Caselli & Sabbadini, 1993).

5.8. Previous studies on the acquisition of articles in bilingual children

Although to date there is a considerable body of research on the acquisition of articles by monolingual children, especially English-speaking ones, there are only a handful of studies investigating the issue in bilingual children, all of which are based on naturalistic rather than experimental data.

Koehn (1994) and Müller (1994) present naturalistic longitudinal data on the acquisition of number and gender in the DP from two French-German bilingual children: Ivar and Caroline. Koehn reports that by age 2;2.21 Ivar must have recognised the notion of singularity vs. plurality for in both languages he uses the German numerals *zwei, "two"*, and the French numeral *deux, "two"*, exclusively to refer to two or more objects. At the same time the indefinite articles *ein* and *un* are only used in conjunction with a singular referent. Number specification therefore seems by and large accurate, while, at the same time, nominals are not correctly inflected for plural or singular. A discrepancy is noted between the semantic notion of number (number specification) and the morphosyntactic agreement between article and nominal. Koehn reports that the grammatical feature number is not available in Ivar's data until 2;6, and it is around this age that in French the child starts making a distinction between the singular definite article *le* and its plural counterpart *les*. Plural

nouns are now systematically accompanied by the appropriate plural article. In German, plural nouns appear, and begin to be used contrastively although not always with the appropriate plural articles.

Koehn points out a difference in the grammatical encoding of plurality in the two languages; in French, Ivar begins by marking number on articles, while in German he starts by acquiring plural noun suffixes first. The implication from these two different acquisitional strategies seems to be that unifunctionality (one-to-one correspondences), and as a consequence consistency and reliability, contributes to facilitate acquisition. Note that, apart from nouns ending in *-al* and *-ail* whose plural suffix is *-aux,* in French the plural is formed by adding */s/* to the singular noun, a phonetically undetectable process[7]. Therefore, from a perceptual point of view Number is clearly marked only on definite articles in French (*le/la,* vs. *les*), and it is exactly on articles that Ivar first marks number distinctions. By contrast, in German articles are not unifunctional, also because of the interaction of the Case feature. For example, *die* is the nominative and accusative form of the definite singular feminine article and also the nominative and accusative form of the plural article. The other forms of the articles, *der, das, des, dem* and *den,* have also homonyms[8]. Given the ambiguity of the German article system it is only natural that the child would start out by focussing on a more reliable paradigm such as the one for noun plural formation. Following Köpcke (1988), Koehn (1994) proposes a possible continuum of schemas for plural nouns on the basis of cue strength which is determined by perceptual salience, type frequency and reliability:

singular				plural	
*	*	*		*	*
monosyllabic	polysyllabic final / r/	polysyllabic final / /	polysyllabic final /s/		polysyllabic final / n/

[7] There are of course exceptions such as a number of suppletive plural forms: *oeil/yeux, ciel/cieux, etc.* , in addition to exceptions: *chandail/chandails, rail/rails, détail/détails, etc.*

[8] Declension of the definite article in German:

Case	singular masculine	singular feminie	singular neuter	plural
nominative	*der*	*die*	*das*	*die*
accusative	*den*	*die*	*das*	*die*
dative	*dem*	*der*	*dem*	*den*
genitive	*des*	*der*	*des*	*der*

On the basis of this schemata the prediction is that the child will be more likely to treat monosyllabic words as singular forms and polysyllabic words ending in / n/, /s/, / / and / r/ as plural forms in this descending order. In German cue strength for the feature number is higher on nouns while in French it is higher on the article system, we therefore expect that the child will start out by concentrating on the system whose cue strength is higher, as is shown by Koehn's data.

In sum the underlying semantic concept of singularity vs. plurality precedes the emergence of the grammatical feature Number, and those elements of the input with higher cue strength will be the ones on which children will initially focus their attention.

Müller (1994) reports that initially number and gender agreement between the determiner and the noun in the DP is only realised in those nominals which contain a definite article, since indefinite articles are analysed as being numerals. The reanalysis of indefinites as articles in their own right is thought to be the trigger of the emergence of the adult grammatical features number and gender.

Berkele (1983) cited in Müller (1994), proposes that definite articles initially encode the [+ definite] or [+S] feature, while [-S] or [- definite] is marked by a null article (cf. Bickerton, 1981, and Cziko, 1982, for a similar claim). The gradual emergence of the grammatical features number and gender correlates with the progressive inclusion of indefinite articles in the children's article system.

Another study of syntactic bilingual acquisition which includes an analysis of article acquisition is Paradis & Genesee (1997). The data comes from a longitudinal study of two English-French bilingual children, Yann and Mathieu, between the ages of 1;9 and 3;0. The authors supply data on the frequency and the productivity of determiner use in the children's two languages and remark that "in contrast to the use of IP clauses, the use of DPs did not seem to be governed by language-specific factors" (Paradis & Genesee, 1997: 114). It is however unfortunate that in the determiner category they collapse definite and indefinite articles, possessives and numerals. Aside from two charts plotting the production of determiners in obligatory contexts in the two languages for each child, there are only a handful of representative examples of the kind of DPs the children produced at different ages. Similarly, there is no indication as to what extent formal features such as gender, number and referentiality were actually part of the children's grammar at any given point.

A very meticulous analysis of a bilingual's nominals is provided by de Houwer (1990) in her study of Kate, an English-Dutch bilingual child. The focus of de Houwer's analysis is essentially on the acquisition of the number and gender features in nominals. For this purpose not only articles, but also demonstratives, pronouns and nouns' plural formation are investigated to give a very informative picture of the child's command of gender and number between 2;7 and 3;4 years of age. Although overall it seems that Kate has figured out the two-way neuter vs. non-neuter gender distinction in Dutch and uses definite articles and demonstratives appropriately 85-88% of the time (the indefinite article is not marked for gender and is therefore not relevant here), her command of all the gender-marking forms is still not complete, and some nouns are still miscategorised. Moreover, from the beginning of the study Kate shows a clear-cut separation between English and Dutch in her treatment of gender. Any errors she makes are fully language-specific and she treats the two languages separately; there is no attempt on her part to apply knowledge of the gender sytem in one language to the other or viceversa. As for Number de Houwer only reports data on nouns' plural formation and not on the determiners themselves.

None of the previous studies on the acquisition of articles in cases of BFLA has investigated what is bilingual children's path to the mastery of the semantic and pragmatic knowledge required to use definite and indefinite articles appropriately. In the analysis of C.'s nominals in chapter 6 we will illustrate the child's performance both on the suppliance of articles in obligatory contexts in general, and also on his mastery of the specific/non-specific distinction and his command of the familiarity condition.

5.9. A bilingual child's task in the acquisition of the English and Italian article systems

As mentioned in section 5.1., DPs encode two distinct sets of nominal features: the interpretable referentiality feature checked by a functional D projection, and non-interpretable phi features ad Case, checked by an AGRN projection below D. Not all languages express all of the above features on the article system. For example, English only encodes referentiality, and Italian encodes gender and number in addition to referentiality, but neither language expresses Case on articles. In English only three articles exist: the definite article *the*, the indefinite article *an* before vowels and *a* elsewhere, and *some* which serves the function of a plural indefinite article. The English definite article is plurifunctional as it serves the function of singular and plural markers,

grammatical gender being non-existent in English, while *a/an's* use must be restricted to singular contexts. Plural generics and mass nouns in English occur without a lexical determiner.

In Italian, articles are subdivided in definite and indefinite, masculine and feminine and singular and plural. Table IV and V below sum up the Italian definite and indefinite article systems respectively:

number	masc.	masc.	fem.
sing.	il	lo	la
plur.	i	gli	le

Table IV. The Italian definite article system

number	masc.	masc.	fem.
sing.	un	uno	una
plur.	dei	degli	delle

Table V. The Italian indefinite article system

Note that there are two singular forms and two plural forms of the masculine article. All simple-consonant-initial words require *il (i* for the plural); words beginning with *s*+ consonant, *z*, *ps* and *gn* require *lo (gli* for the plural).

In Italian articles agree in number and gender with the nouns and adjectives they modify. From a brief overview of the Italian article system it could be argued that the Italian-speaking child will have to acquire three features (referentiality, number and gender) as opposed to one in English (referentiality). Number is not expressed on the English definite article, by contrast the English-acquiring child must realise that only *the* but not *a/an* can be used with singular and plural count nouns. The child must figure out that while *the* is neutral respect to the feature number *a* and *an* are specified as [+singular] and *some* as [-singular]. Also the child must figure out the distinction between count nouns and mass nouns, the latter do not take an article in English. Finally, some phonological analysis must be performed on the basis of the complementary distribution of the two allomorphs of the indefinite article so that *an* will be used before vowels and *a* before consonants. It is reasonable to assume that the segmentation of *a+noun* will isolate the first element as the determiner, which will probably be assumed by the child to be the default form of the indefinite article. This

may lead to mysanalyses of strings *an+noun* beginning with a vowel as *a+(n)noun*, e.g. *an orange* misanalysed as *a norange.*

In addition to acquiring the various features associated with articles, bilingual English-Italian children must also figure out the distribution of articles in the two languages. The child must know *if* and *when* to use an article. Both in Italian and in English some contexts do not allow an overt determiner and instead require a null one. For Italian Longobardi (1994) reports that in a lexically-governed position, plural count nouns, singular mass nouns and singular count nouns in the scope of sentential negation are allowed to appear without an overt determiner and an existential reading is intended. If an overt determiner is present then the generic reading is called for. By contrast, English never tolerates the use of the definite article with plural generics, mass generics or singular proper names. Note however that the definite article is allowed with plural nouns and singular mass nouns in English but the interpretation can only be the specific definite reading. The main difference between English and Italian as far as the distributuion of articles is concerned lies in the restriction of the definite article with plural and mass generics. While in Italian the definite article is always required both for the generic and the specific definite reading, in English it is only required for the latter and impossible for the former. Longobardi (1994) refers to the use of the Italian definite article with generics as *expletive* and observes that "this peculiarity that English displays in limiting the use of the expletive articles as described [i.e. no overt definite articles with proper names, plural or mass generics] might significantly correlate with the lack of morphophonological expression of gender and number on the article" (Longobardi, 1994: 654).

5.10. Predictions for the acquisition of DPs

The aim of the present study is to investigate in a more exhaustive manner the full range of formal, distributional and semantic-pragmatic skills that a child needs to achieve command of the article system. On the basis of the syntax and semantics of DPs illustrated in the previous sections, we are now in a position to formulate a number of predictions with respect to the acquisition of DPs:

(9)a initially, when children are not yet aware of the relevance of either agreement or referentiality features, they will produce nominals without a determiner

(9)b the initial determinerless phase is followed by a stage in which children start to produce their first Article+N combinations, they are almost exclusively singular and lexically restricted to a small number of nouns

(9)c gradually the number of nouns articles appear with increases, and children start to produce Article+N combinations more consistently. By this time there is still no clear evidence that articles are identified as markers of referentiality, but the local head-to-head relation between articles and nouns is appropriately marked by agreement

(9)d the final stage is represented by the identification of the referentiality feature. Not only do children obligatorily satisfy the agreement relations between articles and nouns, they also use definite and indefinite articles to identify referents appropriately

The predictions in (6) envisage an acquisitional sequence where -Interpretable phi features encoded in a local agreement relationship between the article and the noun emerge earlier than the Interpretable referentiality feature. The claim being made here is that the ability to encode referentiality through the appropriate selection of a definite or indefinite article requires that the child be aware of the entity's referentiality; the child must also know whether its referent is [+specific] or [-specific], and in chossing the appropriate definite or indefinite article to mark specificity, the child must be in a position to take into account his/her listener's point of view. Contrary to what initially claimed by Piaget (1962) and Bruner (1966), a number of studies have shown that children are indeed able to make the relevant distinction between specific and non-specific referents from very early on (see Maratsos, 1976; Warden, 1976; Karmiloff-Smith, 1979; Garton, 1983). What children seem to be having more difficulties with is the pragmatic requirement of familiarity.

In sum, the use of articles is cognitively demanding, it requires semantic-pragmatic competence about both the specific/non-specific distinction and the given/new distinction. By contrast, the correct marking of gender and number simply requires the child to register the mapping of these features in a local agreement relationship via distributional learning. It is therefore reasonable to predict that -Interpretable agreement features will be acquired relatively early via low-level

distributional learning, while the Interpretable referentiality feature will require the development of higher level semantic-pragmatic competence.[9]

[9]See López-Ornat (1997) for a similar claim on the acquisition of the morphophonology of articles in child Spanish.

Chapter 6

The acquisition of the article system: semantics, syntax and pragmatics

6.1. Introduction

A number of predictions were made in section 5.10 for the acquisition of the article system and the emergence of interpretable and non-interpretable formal features associated with Determiners. In this chapter we will analyse C.'s use of definite and indefinite articles in Italian and English in order to establish to what extent the child can be credited with nominal functional categories such as AGRN and D. Section 6.2. looks at C.'s production of articles in obligatory contexts in Italian, and at the distribution of definite and indefinite articles in the four contexts identified in section 5.4 as [-S;-F], [+S;-F], [+S;+F], and [-S;+F]. Section 6.3 examines the emergence of gender and number features in Italian. Parallel to section 6.2., section 6.4 looks at the distribution of definite and indefinite articles in English obligatory contexts with respect to different conditions of specificity and familiarity. Section 6.5 is on the emergence of number in English and sections 6.6 and 6.7 compare and contrast the emergence of formal non-interpretable features between the two languages, while section 6.8 reevaluates the data in light of the predictions for the acquisition of DPs in section 5.10.

6.2. Distribution of Italian articles

Table I below reports the number of definite and indefinite articles and the percentages in obligatory contexts across the 20 Italian files:

File	omissions	def. article	indef. article	null article	other	% in o.c.
1	72	-	1	-	-	1.36
2	102	1	-	1	-	1.96
3	80	16	3	-	-	21.34
4	37	13	1	-	3	31.48
5	22	-	-	6	-	21.42
6	101	12	-	1	2	12.06
7	93	27	6	15	-	34.04
8	5	17	24	-	-	89.13
9	29	27	7	-	1	54.68
10	6	80	13	3	-	94.11
11	2	47	10	2	-	96.72
12	3	18	6	3	-	90
13	4	72	11	5	1	95.69
14	3	90	35	5	13	97.94
15	-	55	19	3	8	100
16	4	13	6	1	2	84.61
17	4	52	12	19	15	87.5
18	1	39	25	6	2	98.63
19	4	30	5	10	4	92.45
20	1	83	29	3	9	99.2

Table I. The number of definite, indefinite and zero articles across the 20 Italian files

The heading "null article" in the fifth column indicates the number of correct 0 article options with mass nouns and plural generics in existential constructions, family terms and vocatives. "Other" in the sixth column indicates the number of all other determiners collapsed together (e.g. demonstrative adjectives, numerals, etc.).

Production of articles in obligatory contexts, and of determiners more in general, does not reach a significant proportion until file 8 (2;1.23), after which it hovers between 84.61% and 100% with a low of 54.68% in file 9 where however 27 out of the 29 missing articles are found in combination with only one noun *afante* (= *elefante, "elephant")* in the construction *dè afante?(= dov'è l'elefante, "where is the elephant?").*

Let us know take a closer look at the distribution of definite and indefinite articles in the four contexts identified as [-S;-F]; [+S;-F]; [+S;+F] and [-S;+F] in section 5.4..

In the [-S;-F] context I included all instances of naming as in *Questo/a è un/uno/una x, "This is a x".* I follow Maratsos (1976) in considering naming a purely taxonomic operation where the DP in predicative position is [-referential] since the referent is seen as any one member of a class rather than as a specific instance of a particular member of a class.

File	un	uno	una	null article	%of naming	errors
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	2	-	-	-	100	-
4	1	-	-	-	100	-
5	-	-	-	4	100	-
6	-	-	-	1	100	-
7	4	-	-	10	100	-
8	13	-	5	-	100	-
9	5	-	1	-	50	-
10	8	-	3	3	100	-
11	-	1	2	2	80	-
12	3	-	1	3	100	-
13	4	-	2	4	80	-
14	10	-	5	4	78	-
15	12	-	2	1	40	-
16	-	-	1	1	0	-
17	2	-	2	7	63	-
18	19	-	4	3	38	-
19	3	-	1	6	60	-
20	5	-	-	-	100	-

Table II. Distribution of indefinite and null articles in [-S;-F] contexts

As Table II clearly shows, indefinite articles in [-S;-F] contexts are exclusively used for naming when articles first appear until file 11 (2;3.7) where the proportion decreases to 80% to go down as low as 38% in file 18 (2;11.12), an indication that Carlo is starting to use indefinite articles in [-S;-F] contexts other than simply naming. There are no instances of indefinite articles being used inappropriately instead of a definite or of a null article. The null articles in the fifth column refer to bare plurals or mass nouns in naming constructions such as *Questi sono fiori , "These are flowers", Questa è carta , "This is paper"*, where an existential interpretation is intended.

Naming typically calls for an indefinite in structures of the type *Quello è un leone, "That's a lion"*, however there are cases where a definite is used in the same construction *Quello è x, "That is x"*; for example when the child is singling out a specifc definite referent and not simply labelling one. So when looking at a picture of a lion, well-known to both the child and his adult listener, and the child says *Quello è il leone, "That's the lion"*, we do not have an instance of naming although the frame of the presentational construction is the same. What the two types of sentences have in common is instead the same syntactic predicative position. Figure 1 charts the number of DPs in predicative position as opposed to DPs in argument position: subject, object and indirect object from file 8 (age), the file where production of articles reaches a significant proportion for the first time.

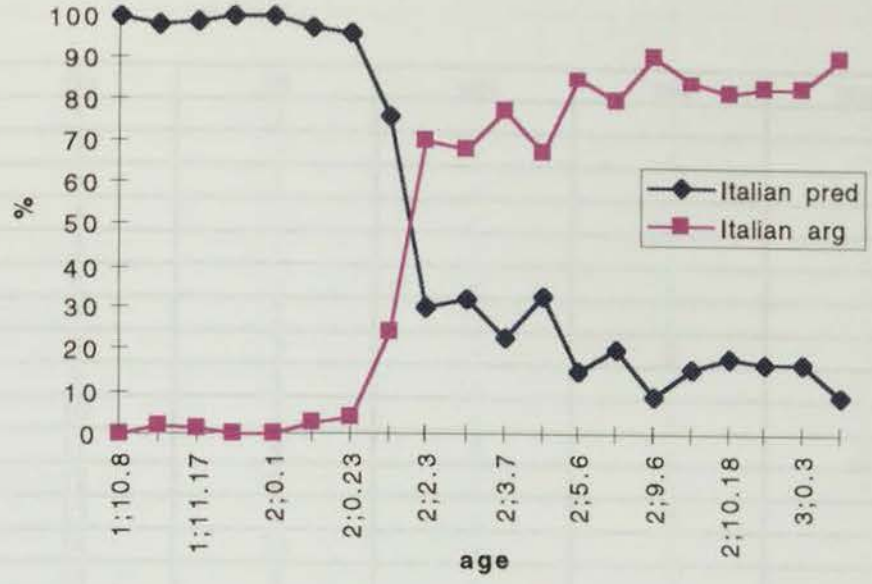


Figure1. The proportion of predicate and argument DPs in Italian

As Figure 1 clearly shows, until file 9 (2;2.3) the proportion of nominals in predicative position is significantly higher than the proportion of nominals in argument position.[1] From file 9 onwards the trend is abruptly reversed and the proportion of referential nominals in argument position becomes significantly larger than non-referential nominals in postcopular position.[2] Two months after his second birthday C. starts to show an interest in the world around him that goes beyond mere labelling. The referents in his discourse gradually change from simple predicates to agents and patients. In a non trivial way this shift correlates with the emergence of verbs in Italian: the use of verbs other than *be* in copular constructions clearly requires the use of arguments, although subjects can remain unexpressed in Italian unless otherwise required by the context (see chapter 8 and chapter 9 for a detailed discussion of verbal predicates).

In Table III data on the use of indefinite articles in [+S;-F] contexts are provided. I included in this context all instances of presentational constructions such as *C'e un/uno/una x, "There is a x"*, unlike the naming contexts where I consider the speaker's intention to be purely of a classificatory nature, in these presentational

[1]The result of a Wilcoxon matched pairs signed ranks test (two-tailed) show $p < .025$.
[2]The result of a Wilcoxon matched pairs signed ranks test (two-tailed) show $p < 0.25$.

constructions there is a clearer deictic intent in singling out a particular member of a class rather than just any member of a class.

file	un	uno	una	errors
1	1	-	-	-
2	-	-	-	-
3	1	-	-	-
4	-	-	-	-
5	-	-	-	-
6	-	-	-	-
7	2	-	-	-
8	3	-	2	-
9	1	-	-	-
10	2	-	-	-
11	4	3	-	1
12	2	-	-	-
13	2	-	3	-
14	16	-	4	-
15	4	-	1	-
16	5	-	-	-
17	5	-	3	-
18	1	-	1	-
19	1	-	-	-
20	10	-	14	-

Table III. Distribution of indefinite articles in [+S;-F] context

Again, indefinite articles are always used appropriately in the [+S;-F] context except for one occasion in file 11 (2;3.7) where a definite should have been used instead, an example of incoherence error. Although Table III above shows that C. correctly produced a number of indefinite articles that are appropriate in a [+S;-F] context, nevertheless the possibility still remains that he might also be producing definite articles which are inherently [+S;+F], except in generic contexts where they are [-S;+F], in contexts which are [+S;-F] thus making those egocentric errors that have so often been reported in the literature. In Table IV the overall number of definite articles is reported for each file with any egocentric errors that C. has committed. Articulated prepositions are also included in the count since in Italian the definite article cliticizes to the preposition.

File	il	lo	la	i	gli	le	errors
1	-	-	-	-	-	-	-
2	1	-	-	-	-	-	-
3	4	11	1	-	-	-	-
4	5	5	3	-	-	-	-
5	-	-	-	-	-	-	-
6	3	5	4	-	-	-	-
7	9	3	12	2	-	1	-
8	7	-	9	1	-	-	-
9	9	7	9	2	-	-	-
10	34	5	34	1	-	6	2
11	20	4	17	2	-	4	1
12	6	1	10	-	-	1	2
13	31	-	30	4	-	7	-
14	49	2	25	5	1	8	8
15	28	2	19	-	-	-	2
16	8	-	1	2	-	2	3
17	20	3	20	4	2	3	2
18	9	1	20	4	1	4	4
19	13	2	6	3	-	2	-
20	21	7	33	5	3	11	1

Table IV. The distribution of definite articles in [+S;+F} contexts

Not unexpectedly a number of egocentric errors appear where an indefinite or null article should have been used instead of the definite one since the listener had no way to identify the referent as specific either through previous mention or through the situational context. It must be noted that in a number of cases where C. uses the definite article an adult may have used an indefinite instead. For example when looking at picture book C. may say something like *C'è il topolino, "There's the little mouse"*, or *Quello è il leoncino ,"That's the little lion"*, where an adult looking at the same picture might have been more inclined to use an indefinite. However previous contact with the book together with his adult interlocutor is sufficient evidence for C. to identify that mouse or that lion as a particular one which is supposedly also familiar to his adult partner. The number of definite articles is therefore largely biased by the previous shared knowledge he has with his listener and by the deictic bias of the here and now of the situational context.

The last identified context for article use is the one for generics [-S;+F]. Not surprisingly there are only a handful of generic DPs starting from file 15 (2;8.4) shown in Table V:

file	il	lo	la	i	gli	le	*0
15	2	-	1	1	-	2	-
16	-	-	-	-	-	-	-
17	-	-	-	-	-	-	-
18	-	-	-	-	-	-	-
19	-	-	1	1	1	1	3
20	-	-	-	1	1	1	-

Table V. Distribution of definite articles in generic DPs [-S;+F]

These are all contexts where it is reasonable to assume that the child is indeed talking about referents that are not obviously deictically identifiable, where therefore not the specific definite reading is intended but the generic one. Of course there is still the chance that, although he seems to be using these definite articles in the generic reading as an adult would do, he may still be having a specific definite referent in mind. Unfortunately because sometimes some of these referents are also physically present during the conversation it is not always possible to disambiguate between the two contexts but the tokens in Table V are potential instances of generic DPs. Examples of generic nominals are given below:

(1) File 15

*CAR: mangiano l'erba le mucche.
%eng: cows eat grass.

(2) File 15

*LUD: per esempio ti piace il pollo?
%eng: for example do you like chicken?
* CAR: sí.
%eng: yes.
*LUD: sí.
%eng: yes.
*CAR: mi piace il pollo.

In addition to some apparently target-like exapmles of generic nominals there are also some incorrect determinerless examples:

(3) File 19

*CAR: cosa mangiano giraffe?
%eng: what do giraffes eat?
*LUD: credo che mangino l'erba.
%eng: I think they eat grass.

(4) File 19

*CAR: no scimmiette mangiano gli alberi
%eng: no monkeys eat trees.
*LUD: e le tigri mangiano +//.
%eng: and tigers eat +//.

(5) File 19

*LUD: e che cosa ti piace mangiare?
%eng: and what do you like to eat?
*LUD: dimmi un po'.
%eng: tell me.
*CAR: patatine mi piace.
%eng: chips I like.

In examples (3) (4) and (5) above C. incorrectly omits the article with what would be considered to be generic nominals in adult Italian. It could be that he simply omits the article because of residual optionality in the use of articles in general, or it could be the result of an interference from English where plural generics do not take a definite article. Given the paucity of data it is not possible to decide between these two alternatives, more data on C.'s use of generics in both languages would be needed to establish whether we are dealing with a systematic crosslinguistic interference, or whether it is simply the result of residual optionality in Italian. In Figure 2 the number of correct definite and indefinite articles in the five contexts are represented graphically:

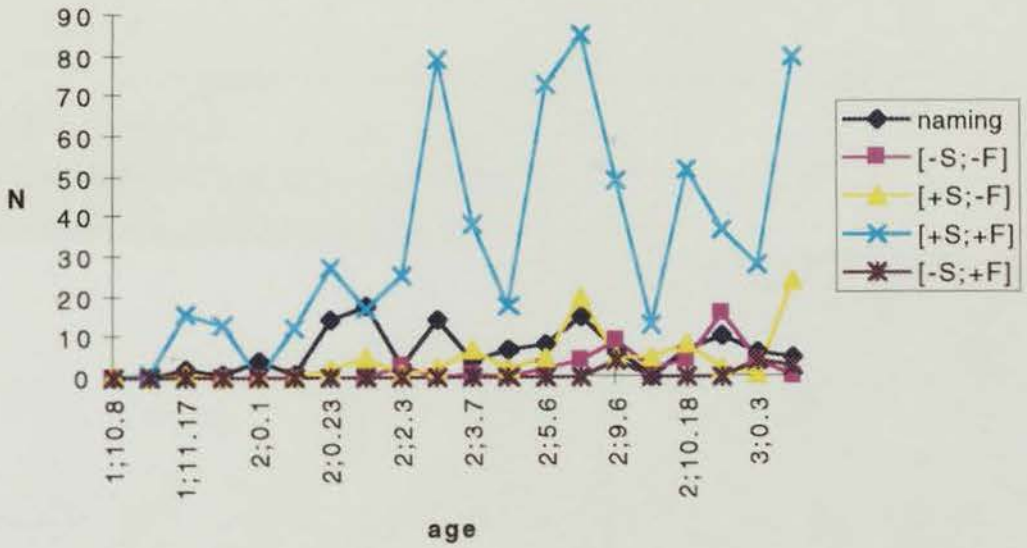


Figure 2. The number of correct definite and indefinite Italian articles in the different semantic-pragmatic contexts

The single most striking fact emerging from the data plotted in Figure 2 is the large number of definite articles in [+S;+F] contexts (N = 651) overall, and when compared to the number of definite and indefinite articles in other contexts. Until the consistent emergence of verbal predicates other than the copula, around file 13 (2;5.6), the majority of definite articles are used in presentational constructions of the type *Questo è il coniglietto, "This is the bunny"*. From file 13 onwards, the inceasing number of utterances containing a lexical verb goes hand in hand with a growing number of definite articles used in referential argument DPs. There are therefore two different uses of the definite article in [+S;+F] contexts: the use in presentational constructions, which is more frequent until the emergence of verbal predicates, and the use in argument DPs, which is related to the use of verbs.

As already mentioned above, not unexpectedly there are only a few examples of definite articles in [-S;+F] environments used with generics (N = 13), and even in those cases there is no certainty that what would be considered a DP with a generic interpretation in adult language is treated in the same way by the child.

As for indefinite articles the largest overall number is found in naming contexts (N = 126), followed by [+S;-F] (N= 90) and by other [-S;-F] contexts (N = 45). The use of indefinite articles in [+S;-F] contexts appears in file 11 (2;3.7), when C. begins to produce verbal predicates in a more consistent fashion.

6.3. Number and gender agreement features in Italian nominals

As far as number and gender are concerned the data only contain a very small number of errors (N = 15); only two of which are systematic and interestingly reflect the overgeneralization of a morphological paradigm while the rest may well be accounted for as performance errors given the very high number of correct tokens.

file	N of plural DPs	% of plural DPs	N of errors
1	-	-	-
2	1	50	2
3	-	-	-
4	3	17.6	-
5	4	66.6	-
6	3	20	1
7	3	6.5	-
8	1	2.4	-
9	3	8.5	-
10	7	7.2	1
11	6	10.1	2
12	1	3.7	-
13	11	12.3	-
14	17	11.8	4
15	3	3.5	-
16	4	18.1	-
17	20	20.4	-
18	9	10.1	2
19	9	18.3	-
20	25	20.16	3

Table VI. Number and % of plural DPs and number of errors in Italian

Aside from file 2 and file 5 where plural DPs account for 50% and 66.6% of the total number of DPs respectively, the proportion of plurals ranges from 2.4 (file 8) to 20.16 (file 20)[3]. The notion of plurality is present as early as file 4 (1;11.25) and it is encoded on the noun where the plural ending is correctly supplied and on the determiner, the numeral *due* (*two*) which is initially used to indicate any set of objects larger than one.

(6) File 4

*LUD: fanne un' altra.
%eng: make another one.
*CAR: due rane.
%eng: two frogs.

(7)
LUD: mmhm quante giraffe ci sono?
%eng: how many giraffes are there?
*CAR: due.
%eng: two.
*LUD: no, una, due e tre!
%eng: no, one, two and three!

[3]Note however that in both files the absolute number of tokens is extremely low (2 for file 2, and 6 for file 5) and therefore the relatively percentages lose most of their statistical significance.

From file 7 (2;0.23) onwards not only numerals but also articles and other determiners (mainly the demonstratives *questi, "these",* and *quelli, "those"*) appear in plural DPs.

The fourth column in Table VI indicates the number of errors found across the 20 Italian files: specification errors (4), agreement errors (1), and morphological errors (10). The four specification errors involve 3 number specification errors where C. uses a singular noun instead of the required plural; and one gender specification error where instead of the feminine he supplies the masculine. As for the only agreement error in the whole corpus it involves lack of number agreement between a plural demonstrative (*questi, "these"*) and the selected noun (the singular *libro, "book",* instead of the plural *libri, "books"*), and note that C. immediately corrects himself by using the appropriate singular determiner in the following utterance:

(8) File 14

*CAR: tu me li prendi.
%eng: you get them for me.
*LUD: cosa faccio io?
%eng: what do I do?
*CAR: anche questi libro.
%eng: these book too.
*CAR: anche questo libro.
%eng: this book too.
*LUD: sí.
%eng: yes.

Given the extremely low number of specification and agreement errors we can safely account for them in terms of performance errors. C.'s performance is even more impressive in the singular where all nouns agree in number and gender with their determiner. There are however a number of segmentation errors where the allomorph *l'* of the singular masculine *lo* is taken to be part of the noun and therefore another determiner appears before the article+noun compound such as in *il l'elefante, "the the elephant",* (file 20, 3;0.17) and *quello l'orsacchiotto, "that the bear"* (file 12, 2;4.14).

A small number of morphological errors are also present in the corpus and are particularly interesting in that they reflect some underlying morphological generalization at work. Three involve misselection of the plural masculine article *i* and *l'* instead of *gli.* This is not surprising since *gli, / /* (the plural form of *lo*) appears in phonetically restricted contexts and is therefore the marked form of the masculine

article. Moreover the lateral approximant in *gli* is certainly not one of the easiest sounds from an articulatory point of view.

The other seven morphological errors involve two irregular plurals (*l' uovo/le uova = egg/eggs,* and *il corno/le corna = horn/horns)* and a singular feminine noun ending in *-e* rather than in the default *-a* singular feminine marker (*la tigre*). In file 10 (2;2.17) when asked *Cosa sono queste?* (*What are these?*) while pointing to a picture of eggs, C. answers *uove* instead of the correct form *uova.* In file 14 (2;5.26) C. produces *la uova, una uova* and *le uove* clearly showing a morphological regularization process whereby the singular has changed from masculine (*uovo)* to feminine *(uova)* and the plural is formed accordingly as with any other feminine noun ending in *-a* in the singular.

The same analogic process occurs with *il corno/le corna* in file 18 (2;11.12) where the plural becomes *le corne* instead of *le corna.* C.'s regularization process is most certainly his solution to what he perceives to be a conflict of competing cues. On the one hand he hears *le uova* in the adult input referring to a plural referent, what is more the plural article *le* also suggests that it must be a plural referent. These semantic and mophosyntactic cues are however in conflict with the *-a* ending on the noun which is a typically singular feminine ending, not a plural one. The semantic cue and the morphosyntactic cue on the article win out and the noun is regularized into the plural *uove,* whose singular form is then *uova.* The same applies to *le corne.* One might wonder why C. has not chosen the singular masculine form *l'uovo* and created the plural *gli uovi* or *i uovi.* The reason is probably the higher frequency in the input of the plural *le uova* as opposed to the singular *l'uovo.* Certainly in the session where these errors were recorded the context (baking a cake) suggests a preference for the noun as being used as a plural ingredient. This is even more plausible in the case of *le corna* vs. *il corno.* When describing an animal, for instance a deer, one is more likely to say something like *Queste sono le corna (Theses are the horns)* or *Che corna lunghe (What long horns)* rather than using the singular. It is thus probable that C. might have only heard the plural *corna* and proceeded to a regularization process whereby *corna* is the singular and *corne* the plural, much in the same way as he does for *uova/uove.* The frequency effect could also account for the fact that other more common and frequent irregular plurals, such as for example *la mano/le mani* (*the hand/the hands)* and *il dito/le dita,* are used correctly throughout the period of observation.

Another cue competition resulting in a morphological overregularization is shown in the extract below from file 20:

(9) File 20

*CAR: vuoi la storia della tigra?
%cmg: do you want the story of the tiger?
*CAR: la [/-] vuoi la storia <delle tigr> [//] della tigr +...
%eng: do you want the story of the of the +...
*LUD: tigre.
%eng: tiger.
*CAR: della tigre?
%eng: of the tiger?
*LUD: sì grazie.
%eng: yes thank you.

The noun *tigre*, *"tiger"*, is feminine and is one of a number of nouns ending in the *-e* suffix which in itself is neither typically feminine nor typically masculine unlike the default feminine-*a* suffix and the default masculine *-o* suffix. The article is feminine, a strong morphosyntactic cue that the following noun must be feminine as well if it is to agree with the determiner. A search through the mental lexicon provides either the underspecified entry *tigr* (see the false starts and the indecision in line 351) or the target *tigre* which is however not a prototypical feminine noun given the *-e* ending. In the cue competition that ensues the morphosyntactic cue on the article overrides its other competitors and gives rise to the overregularization of the underspecifed *tigre* into the canonically feminine *tigra*.

There is good evidence in C.'s data that morphosyntactic number and gender agreement in the DP are well-established from the earliest stages of determiner production. In fact the strong morphosyntactic cue provided by the article overrides any other type of information when it comes to choosing the number and the gender of the noun modified by it.

6.4. Distribution of English definite and indefinite articles

Along similar lines to the previous section, in this second part we will focus on C.'s performance in the use of English articles.

The overall number of articles and other determiners and the percentage in obligatory contexts are reported in Table VII below:

File	omissions	def. article	indef. article	null article	other	% in o.c.
1	15	-	-	-	-	0
2	24	-	-	-	-	0
3	37	1	6	-	-	15.9
4	19	-	-	-	-	0
5	11	-	-	-	-	0
6	18	2	-	-	-	10
7	8	2	11	-	-	61.9
8	16	3	29	-	1	67.34
9	12	11	12	1	5	70
10	-	1	12	2	2	100
11	4	2	15	3	8	87.5
12	2	25	17	9	20	97.26
13	6	15	59	1	8	92.77
14	4	19	67	2	23	96.52
15	-	16	12	5	3	100
16	2	12	25	-	12	96.07
17	2	18	13	2	6	95.12
18	1	28	15	3	6	98.11
19	3	25	18	1	7	94.44

Table VII. Distribution of definite and indefinite articles in English

The heading "null article" in the fifth column indicates the number of correct 0 article options with mass nouns and plural generics in existential constructions, family terms and vocatives. "Other" in the sixth column indicates the number of all other determiners collapsed together (e.g. demonstrative adjectives, numerals, etc.).

Although as early as file 7 (2;2.12) production of articles in obligatory contexts reaches 61.9% and increases to 75% in file 9 (2;4.7) to stabilise around 95% in subsequent files, the raw figures are in a way misleading and the risk is there to overestimate Carlo's actual mastery of the article system by age 2;4.7. In file 7 out of a total of 13 noun phrases articles appear in combination with only four different nouns: 6 *that a dog,* 4 *that a ball,* 2 *where the ball,* 1 *a tiger go uaaah* . A similar pattern is observed in file 8 (2;2.24) and 9 (2;4.7). In file 8 out of 32 noun phrases the three noun phrases *a cat* (6 times), *a dog* (7 times), *a mouse* (5 times) account for eighteen of them; ten remaining noun phrases are *that a head* (3 times), *that a picture* (3 times), *that a duck* (twice), *the ball* (twice), *the pencil* (once), *the cocodile eyes* (once), *a paper* (once). Similarly in file 9 fifteen out of thirty noun articles are in combination with only three different nouns: *a piece* (6 times), *the cat* (6 times) and *it's a teddy* (3 times). By contrast in the corresponding Italian files 10 (2;2.17) and 12 (2;4.14) the type/token ratio is much higher (0.382 type/token ration in English file 8 compared with 0.67 type/token ratio in Italian file 10), indicating that there is considerable less repetition of

the same nouns. The small number of different nouns with which Carlo uses articles in English compared to the larger repertoire he draws upon in Italian is indicative of a formulaicity and a lexically bound approach which call for a degree of scepticism in the interpretation of these initial Art +N combinations. There is reason to believe that these articles are not adult-like realisation of determiners, the degree of lexical specificity which characterises their use leads us to believe that there is no sense in which C. has realised the function of articles in an across the board fashion.

In order to assess Carlo's mastery of semantic and pragmatic skills in the use of definite and indefinite articles proportions were computed for the four different environments identified earlier as [-S;-F]; [+S;-F]; [+S;+F] and [-S;+F]. Table VIII below shows the number of indefinite articles used in [-S;-F] contexts including naming:

File	a/an	0	% of naming	errors
1	-	9	100	
2	-	-	-	-
3	4	-	100	-
4	-	-	-	-
5	-	-	-	-
6	-	-	-	-
7	11	-	100	-
8	29	-	100	1(a paper)
9	11	-	100	-
10	9	2	100	-
11	11	3	100	-
12	10	9	100	-
13	26	1	22.22	-
14	19	2	61.90	-
15	11	5	87.5	-
16	11	-	18.18	-
17	7	-	85.71	-
18	13	2	60	
19	14	1	21.42	-

Table VIII. Distribution of English indefinite and null articles in [-S;-F] contexts

Only one incorrect use of the indefinite article with a mass noun occurs: file 8 (*a paper*) where Carlo treats *paper* as a count noun, possibly in an elliptical construction of the type *a (piece of) paper.* Also to be noted that the correct use of a null article with plural and mass nouns in file 1 is most probably not an indication that Carlo has already figured out the mass/count distinction but the reflex of the use of null articles across the board at that age. There is however considerable evidence (from file

12, 2;9.6, onwards) that whenever mass nouns or plural nouns not in the specific definite reading appear, they are never modified by an indefinite article. Surely an indication that a target-like discrimination between mass and count nouns and between existential and specific definite reading is already operational in Carlo's system.

In the English [-S;-F] contexts, as was observed for the Italian [-S;-F] contexts, there is a high proportion of noun phrases in naming contexts. It is not until file 12 (2;9.6) that indefinite and null articles start to be used in [-S;-F] contexts other than naming. Not surprisingly this coincides with a shift from more descriptive tasks such as book reading to more dynamic activities such as building Lego constructions or assembling jigsaw puzzles. As shown in Table XIX below, the number of indefinites in [+S;-F] contexts also gradually increases from file 9 (2;4.7) onwards:

File	a/an	0	errors
1	-	-	-
2	-	-	-
3	2	-	-
4	-	-	-
5	-	-	-
6	-	-	-
7	-	-	-
8	-	-	-
9	7	-	-
10	3	-	-
11	3	-	-
12	4	-	-
13	33	-	-
14	48	-	-
15	1	1	-
16	14	-	-
17	6	-	-
18	2	1	-
19	4	-	-

Table XIX. Distribution of English indefinite and null articles in [+S;-F] contexts

The two peaks in file 13 (2;10.1) and 14 (2;10.15) are once again related to the type of activity Carlo was engaged in in those two particular sessions. In this case it was playing picture domino where Carlo and his adult partner took turns in matching picture cards until one player ran out of cards. This elicited many *I put a x* where x was the animal on the card that he wanted to match. Because these *I put a x* utterances were accompanied by the gesture of selecting a particular card with the named animal I considered them to be instances of [+S;-F] contexts where a specific member of a class is selected and is introduced as a specific referent to the speaker [+S] but still new to the listener [-F]. By contrast, I classified an apparently similar type of utterances such as *I can put a x* as [-S;-F], where Carlo is simply considering the idea of what kind of animal he could look for and put, but he has not yet selected a particular card with the specific animal on it.

The use of indefinite articles predominantly for naming decreases as Carlo's attention gradually shifts from simply labelling the objects around him to actually talking about them.

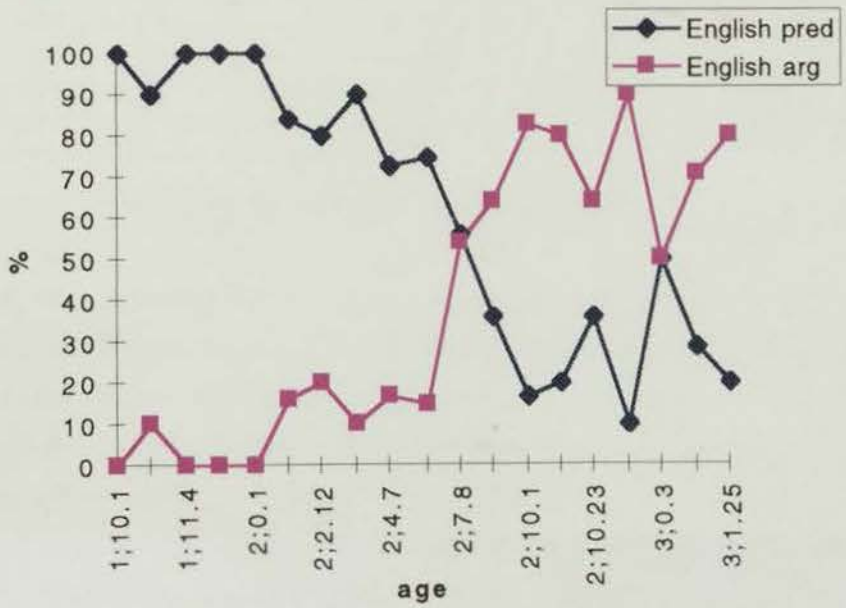


Figure 3. The relative proportion of predicate and argument DPs in English

As shown in figure 3, in file 12 (2;9.6) DPs start appearing in subject and object position rather than in the predicative position of the formulaic *that/this is a x*. Just like in Italian, in English too we observe a transition from the prevalent use of non-referential DPs in predicative position to the use of referential DPs in argument

position. However note that the transition in English occurs much later, around 2;9 compared to 2;2 in Italian.[4]

The distribution of definite articles, typically associated with [+S;+F] contexts is given in Table XX:

File	the	errors
1	-	-
2	-	-
3	1	-
4	-	-
5	-	-
6	2	-
7	2	-
8	3	-
9	11	-
10	3	1
11	4	2
12	25	-
13	15	-
14	19	-
15	16	-
16	12	-
17	18	-
18	27	-
19	25	2

Table XX. Distribution of English definite articles in [+S;+F] contexts

With only five exceptions, Carlo's use of definite articles is appropriate for [+S;+F] contexts. The errors are of the egocentric type where a definite has been used instead of an indefinite in what is really a [+S;-F] context. The small number of errors however does not necessarily imply that Carlo is sensitive to the [±F] feature in the same way an adult would be. Definite articles in a [+S;+F] environment can either be anaphoric or deictic, the deictic use is allowed whenever exophoric reference is intended, i.e. when some extralinguistic information ensures that both speaker and listener share the same knowledge about the referent in question. By contrast, the anaphoric use of definite articles is strictly endophoric and presupposes previous mention of the referent in question. The deictic use is by definition [+F]. Shared knowledge between speaker and listener can safely be assumed in that the universe of discourse is the same for both participants in the conversation; there is no additional

[4]Two separate Wilcoxon matched pairs signed ranks tests were run to test for the significance of the observed difference between referential and non-referential nominals before and after file 12 (2;9.6): $p < .01$ before 2;9.6 and $p < 0.25$ after 2;9.6.

information the speaker needs to take into account. If however the speaker cannot automatically infer that his or her listener shares the same knowledge about the referent in question, the use of a definite article is justified only by previous mention. The use of the definite article in this case is anaphoric and relies on endophoric reference where the introduction of the referent to the universe of discourse has taken place by linguistic rather than extra-linguistic means.

6.5. Number agreement in English DPs

As is well-known, English is not a language that encodes grammatical gender on nouns, determiners and other modifiers, therefore the gender feature is irrelevant in terms of agreement between determiner and noun. Number, on the other hand, is marked in English, but only on nouns and, typically, not on articles and adjectives, although there is a singular/plural distinction in demonstratives such as *that/those, this/these* and quantifiers such as *much/many*. Plurality is thus marked on the noun, typically by the addition of one of the allomorphs of the plural marker */s/*, or through a suppletive form in the case of irregular plurals. Table XXI shows the number of plural DPs, their proportion with respect to the overall number of DPs and any errors.

file	N of plural DPs	% of plural DPs	errors
1	-	-	-
2	-	-	-
3	-	-	-
4	-	-	-
5	-	-	-
6	-	-	-
7	-	-	-
8	2	6	-
9	2	6.8	-
10	-	-	-
11	2	7.1	-
12	6	8.2	-
13	11	13.2	1
14	22	19.8	1
15	10	27.7	1
16	9	18.3	-
17	-	-	-
18	15	28.8	1
19	4	7.8	-

Table 18. Number and proportion of plural DPs and number agreement and specification errors

Although plurals in English appear later than in Italian (2;2.24 as opposed to 1;11.25), the percentages are comparable to the ones found in the Italian data with a range between 6 and 28.8%. What is striking is the almost complete absence of errors, only 3 since the error in file 15 (2;10.23) is a repetition of the one in file 14 (2;10.15): *a stairs,* where C. takes the invariable plural noun *stairs* to be a singular count noun. As for the error in file 13 (2;10.1), it is a specification error where the singular *apple* is supplied instead of the plural *apples* when answering the question *What are these?*[5]

The error in file 18 (3;0.16) is an interesting violation of the morphological constraint that applies to compounding and inflection whereby only suppletive forms can appear inside a compound, and not regularly inflected ones (Kiparsky, 1982; Borer, 1988; Gordon, 1993). In pointing to two tiger cubs C. says *two babies-tiger* instead of the correct *two baby-tigers.* Alternatively one could think of *two babies tiger* as a head noun *(tiger)* modified by *babies* which in this case is not treated by C. as adjectival noun, which is invariable in English, but as an ordinary noun which takes the plural marker *-s.* or perhaps like an Italian adjective that inflects for number. However if this is the case *tiger* should have been equally plural but it isn't . Yet another possibility is to think of *tiger* as a postmodifier.

A variety of determiners are found in C.'s plural DPs: zero articles, possessives *(my, your),* numerals *(two, three)* and to a lesser degree demonstratives such as *these* and *those.* Although admittedly less than half of all determiners are markedly plural (40%) *(26 two, 1 three, 3 these),* agreement is always correct, and the singular indefinite article *a* which is inherently singular is always followed by a singular noun with the exception of the *a stairs* error.

6.6. Similarities and differences in the production of referential DPs in Italian and English

Prima facie, by simply looking at the percentages of definite and indefinite articles in obligatory contexts (Table I and Table VII), one might be inclined to say that articles are being acquired simultaneously in the two languages. In both languages there is an initial phase where articles are either totally absent or produced only sporadically. The occasional appearance of articles in this initial phase may indicate that C. is indeed beginning to notice these functors, although their precise function and

[5] Alternatively one could say that C. supplies *apple* as a mass noun, although it is unlikely in light of the fact that whenever he uses this noun he provides an appropriate article.

morphosyntactic realisation still elude him at this point. No systematicity is obviously operational in the child's grammar yet, the same nouns that occasionally appear with an article also appear without one in the same recording.

The semantic and syntactic features associated with articles and the notion of agreement dependent on the realization of an abstract AGRN projection and a D projection do not play a role in the child's grammar and in his pragmatic competence at this point. In the absence of a functional category, which anchors the sentence to the universe of discourse, the child assumes by default that the referent is specific and familiar, hence there is no need of further specification through the use of an article (Hyams, 1994).

The immature pragmatic competence that allows the child this default specific and familiar interpretation is often greatly facilitated by the situational context in which most adult-child interactions take place in the early years of life. Two key features of these early adult-child interactions are minimally responsible for the child's treatment of referents as familiar: firstly, when talking to young children adults tend to refer to "the here and now". They situate their utterances in the present, rather than in the future or in the past. Also they are more likely to talk about referents that are physically present, and that the child can easily see and possibly touch, rather than about referents that are out of sight and out of reach. In other words deictic reference is central to these early interactional exchanges, both in spatial (here) and temporal (now) terms. Another defining characteristic of early adult-child conversations which is a direct consequence of the spatio-temporal deictic nature of the exchanges, is the participants' shared knowledge of the situation. The physical presence of the participants in the conversation, together with the physical presence of the referents, ensures that all parties involved share the same experience of the discourse context. The physical presence of the participants is as crucial as the physical presence of the referents in their identification as familiar, since one could envisage a situation where a child is talking about referents that are deictically identifiable by him or her but not by the adult (i.e. telephone conversation, calling from another room, etc.)

Given that in most common adult-child conversation settings such as looking at picture books, playing with toys, watching television, etc., shared knowledge can be inferred by the availability of the referents in questions to both participants, treating the referents in question as familiar is a perfectly safe option for the child. Moreover, parents often tend to engage in naming routines with their children where all they are

eliciting is the labelling of an object[6]. In predicative constructions of the type *this is a x*, the DP in predicative position is [-referential], hence the notion of individual member of a class is not applicable here. The indexical function of the article which determines specific definite reference in referential DPs is irrelevant in predicative DPs where the article does not serve such a referential function. It is then perfectly natural that children may initially overlook articles and other determiners, i.e. until they discover that they too have a sepcific indexical function.

If the initial zero article option is justified by the default familiarity assumption, hence a large proportion of determinerless nominals in formulas of the type *That's x*, slowly but gradually the child starts to notice articles and to use them. Thus, alongside *That's x,* we also find *That's a x* and *That's the x.* At first definite and indefinite articles may be in free variation in naming constructions, simple placeholders that have not yet been differentiated according to their different pragmatic and semantic functions. It is when the child starts to grasp the notion of any member of a class as opposed to a particular member of a class that the child starts discriminating between the use of the indefinite and the definite article. The former being used in [-S;-F] environments and the latter in [+S;+F] contexts.

Some support for a default zero option in early child language for naming comes from the use of zero article for naming in creole languages (Bickerton, 1985 quoted in Cziko, 1986). Here too, naming does not require the additional use of an article; Bickerton (1981) also reports that in creole languages a zero article is used with all [-S] referents.

Similarly to monolingual English and Italian children then, C. goes through an initial phase where articles are generally omitted. In English it is around age 2;2 (file 7) that for the first time we find a percentage in obligatory contexts indicating that articles are beginning to emerge. In Italian, like in English, a significant proportion of articles in obligatory contexts start to appear around age 2;2 (file 8), although the proportion is higher than the initial proportion in English, 89.13% vs. 61.9%. If we discount a drop to 54.68% in file 9 (2;2.4) in Italian where 27 out of the 29 missing articles are in the self-repeated question *dè elefante? (= dov'è l'elefante?, "where is the elephant?"),* we see that the production of articles stabilises around 95% from file 10 (2;2.17). A similar picture emerges for English where, after file 9 (2;4.7) percentages remain stable

[6]See Snow (1981) , Bruner (1983) and Peters & Boggs (1986) for the importance of interactional routines in the management of joint attention and reference in language acquisition.

around 95%. A time lag of two months intervenes between the first data point where C. reaches target-like performance in Italian (2;2.17) and in English (2;4.7).

Furthermore, although the proportions of articles supplied in the two languages may be comparable, the actual number of tokens in these initial stages is smaller in English than it is in Italian: 13 for English file 7 (2;2.12) vs. 41 for Italian file 8 (2;1.23) ; 32 for English file 8 (2;2.24) vs. 34 for Italian file 9 (2;2.3); 24 for English file 9 (2;4.7) vs. 59 for Italian file 11 (2;3.7) . If one looks at the actual number of nouns that appear with an article, there is a clear difference between the two languages. C.'s early production of articles in English is extremely stereotyped and formulaic, characterized by numerous repetitions of adult utterances and self-repetitions. Virtually only indefinite articles in naming contexts are present, and it seems that C.'s main preoccupation at this stage is labelling the objects around him. He is not interested in identifying a referent, thus introducing it to the universe of discourse so that he can subsequently refer to it. At least in these initial phases, naming is C.'s primary mode of linguistic interaction in English. This is even more apparent when looking at the number of DPs in predicative position as opposed to DPs in argument position. It is not until file 12 (2;9.6) that for the first time the number of non-referential DPs is lower than the number of referential DPs in argument positions. Until then the majority of DPs appear as predicates in formulaic structures of the type *That's a x*.

By contrast, in Italian the shift from the use of articles purely in a naming function to the identification of referential DPs occurs much earlier. As early as file 9 (2;2.3) the number of referential DPs in argument position is higher than the number of non-referential DPs, a trend that continues throughout the period of observation. There is a delay of seven months in English with respect to the emergence of referential DPs, which coincides with C.'s return from the United States where he had spent the previous six weeks with his family and American relatives. The shift from non-referential to referential DPs clearly ties in with the emergence of verbs which appear earlier in Italian than they do in English (see chapter 9 for a comparison of the emergence of verbal predicates). For instance the large number of object DPs with an indefinite article in the English files 13 (2;10.1) and 14 (2;10.15) correlates with the repeated use of two verbal frames : *I put a x* and *I can put a x* while playing picture domino. The repetitive nature of the game, where C. and his adult partner take turns in matching picture cards, is the main reason for the many self-repetitions of these two frames modelled on the adult input.

Utilizing measures such as the type/token ratio of nouns used with an article, the number of non-referential vs. referential DPs, and the number of indefinite articles used in naming constructions, has allowed us to gain a clearer picture of the data than a purely quantitative analysis of the raw percentages could have done. Other indicators such as number and gender specification and agreement, and the use of definite and indefinite articles in different semantic and pragmatic contexts, can also be used to achieve an even more precise understanding of C.'s competence in the two languages.

6.7. Number and gender agreement

Because English does not encode gender and number in the DP to the same extent that Italian does, a direct parallel between the two languages in terms of agreement is not strictly applicable. Nonetheless there are conclusions that can be drawn by examining C.'s performance in the two languages.

The analysis of the Italian data provides a rather impressive picture of C.'s competence in the mastery of gender and number from very early on.Virtually as soon as determiners appear they invariably agree with the noun they accompany with the few exceptions noted above in section 6.3, which are in themselves highly informative about C.'s approach to nominal morphology. This is also consistent with previous research on the acquisition of Italian morphology: omission errors and not commission errors account for the vast majority of children's initial target-deviant performance[7].

Mastery of the Italian agreement system is indeed some achievement when one considers the amount of information that it encodes. Firstly, the child must realize that nouns fall into different classes which for convenience we label feminine and masculine gender, and which are associated with different types of determiners (feminine and masculine). Membership of one of these two classes is an intrinsic property of the noun which is invariable, i.e. a noun is either masculine or feminine and this is marked by specific morphophonological properties which in Italian correspond to default *-a* suffix for feminine nouns and default *-o* ending for masculine nouns. In addition, the child must deal with a class of nouns ending in *-e* whose membership in the feminine or masculine class is only detectable through the association with articles and adjectives which modify them (*il cane bianco, "the white dog"*, but *la neve bianca, "the white snow"*). Besides learning about gender classes the child must also grasp the notion of

[7]See Chini (1997) for a very informative and exhaustive view of the acquisition of morphological gender in Italian and other languages.

plurality vs. singularity and learn that there is a morphosyntactic correlate of singular and plural which in Italian is marked on noun endings and articles, and is also strictly connected to gender: each gender has its own plural marker, there is not a default one like in English or Spanish.

In sum, the Italian-speaking child must master the article paradigm, and eventually that of other determiners, the nominal inflectional system which comprises two genders and two numbers, and the fundamental notion of agreement inside the DP between nouns and determiners. This might seem like a daunting task but there is good evidence that, as Pizzuto & Caselli conclude in their study on the acquisition of Italian morphology, "gender appears to be learned very quickly in Italian. This is not surprising if we remember that there is a great deal of consistent information for gender categorization in this language" (Pizzuto & Caselli, 1992:551). The same can be said for number marking, which is also intimately connected to gender. We believe that it is in fact the very consistency, redundancy, transparency and availability of morphophonological cues in Italian nominal morphology which makes the learner's task not only possible, but greatly facilitates it.

The central role of language-specific morphophonological cues in the acquisition of morphological paradigms and local morphosyntactic dependencies, such as the one between articles and nouns, has convincingly been shown by a number of researchers working on connectionist approaches in the past ten years or so. For German (MacWhinney et al., 1989; Taraban, McDonald & MacWhinney, 1989), a connectionist simulation on the acquisition of articles has revelead an impressive resemblance to patterns discovered in naturalistic L1 acquisition. Similarly to L1 learners the net went through an initial phase of article omission, displayed a tendency to overgeneralize the feminine article *die*, and committed errors in non-nominative cases and between masculine and neuter. A similar study on French (Sokolik & Smith, 1992) has shown that, like a child, a neural net can predict the gender of a new noun on the basis of its morphophonological characteristics alone, without having to rely on additional semantic or syntactic information (cf. Karmiloff-Smith, 1979, for similar findings in the initial stages of French acquisition).

A connectionist simulation of the acquisition of the article-noun association in Italian (Pizzuto, Caselli & Tangorra, 1991) where a network had been initially trained on 500 article-noun pairs, proved able not only to "learn" the relevant article- noun associations, but also to generalize them to 290 new nouns, with a few generalization

errors of the kind produced by Italian-speaking children (e.g. *la pigiama* instead of *il pigiama*). Also like L1 learners the network initially committed more omissions than commission errors.

What is not clear from the information available about Pizzuto et al' s (1991) simulation is what was the respective role of the article and of the noun suffix in predicting the correct gender. There is some evidence from experimental study and naturalistic observations on French and Spanish (Karmiloff-Smith, 1979; Clark, 1985; Pérez-Pereira, 1991; Montes Giraldo, 1976) that the initial strategy is to rely on the phonological structure of the noun suffix to predict gender, only later do children pay attention to the syntactic information provided by the article.

From the two overregularisation errors found in the corpus (*la uova/le uove , "the egg/the eggs",* instead of *l'uovo/le uova;* and *la corna/le corne, "the horn/the horns",* instead of *il corno/le corna*) C. seems to privilege a syntactic strategy over a morphophonological one from very early on. When the noun suffix cue and the article cue are in conflict, both times he chooses the article cue as predictor of gender, a strategy that normally gives the correct results, except in the case of irregular nouns where the forms are suppletive and cannot therefore be predicted on the basis of regular inflectional paradigms. The early preference for a syntactic strategy in gender assignment, i.e. agreement between the article and the noun, over a purely morphological one, i.e. the noun ending, reveals an early sensitivity to the syntactic function of determiners. The discovery of this agreement relationship between articles, and determiners in general, and nouns is likely to be greatly facilitated in languages such as Spanish and Italian where there are default gender markers. The one-to-one correspondence between an article, for instance *la*, and a noun ending with an *-a* suffix certainly facilitates in making the relationship between a specific functor and a whole class of nouns more transparent, easily detectable and learnable.

As for English, as mentioned above, gender is not grammatically encoded; number is, although not on the definite article. As shown in 6.5, there are only three errors in C.'s English plural DPs, one of which is not an agreement but a specification error. Overall, he seems to be doing as well in English as in Italian. There is however a difference in the age of emergence of plurals: 1;11.25 in Italian as opposed to 2;2.24 in English. Once again there is a delay of approximately three months between the two languages with Italian in the lead and English following behind. There does not seem to be an immediate trade-off between the emergence of a feature in the dominant language

and the emergence of the same feature in the 'weaker' language. Nonetheless, if this is true, it is only partially so. It is fair to say that C. seems to be 'slower' in English and that he does not immediately transfer what he has discovered in Italian to English; however there seems to be some evidence for a more subtle and indirect positive consequence. With respect to English monolinguals C. shows less optionality in the use of the *-s* plural marker; although the grammatical encoding of plurality appears later than in Italian, and in line with other monolingual children, nevertheless it seems to be more 'stable'. In a way it is as if the established notion of agreement in Italian obliquely fed into the notion of agreement in English (Mervis & Johnson, 1991).

6.8.Knowledge of semantics and pragmatics in article usage.

If that is the situation with respect to knowledge of gender and number agreement, what of the semantics and pragmatics that regulate the use of definite vs. indefinite articles? As shown in the results section, the only systematic errors that C. commits are egocentric errors of the type that is widely reported in the acquisition literature. There are 25 of them in Italian (3.7% of the total number of definite articles) and 5 in English (2.7% of the total number of definite articles). The absolute number is higher in Italian, but the overall number of definite articles is also higher in Italian (N= 651) compared to English (N= 180). The only other error in English is one where C. treats a mass noun like *paper* as a count noun in *this is a paper* (file 8). The fact that he treats other mass nouns such as *milk, juice, chocolate, sand* correctly is an indication that the mass/count distinction is part of his grammar, and that he must know that these nouns belong to a special class that does not require an indefinite article in [±S;-F] environments.

In Italian there is an example of incoherence error where an indefinite is used instead of a definite in a [+S;+F] environment. In addition there are three errors in the [-S;+F] environment of generics; on all three occasions the required definite article is omitted leaving an ungrammatical bare noun in Italian. This is probably a transfer from English where it is precisely this kind of generic plural that requires a zero article. The percentage of errors in generics is however only 23% (3 out of 13) which might indicate that there is an emerging use of definite articles in generic contexts in addition to the widespread use of definites in [+S;+F] contexts.

The number of definite articles is the highest in both languages, although there is a considerable difference in absolute terms (651 in Italian and only 180 in English). One interesting difference between the two languages lies in the relative proportion of

definite articles. In Italian there is a major difference between the number of definite articles and indefinite articles in the various contexts, with 126 indefinites in naming contexts accounting for the largest number after definites in [+S;+F] contexts (N = 638). In English, on the other hand, the difference between the number of definites in [+S;+F] environments and the next largest number of articles (indefinites in naming contexts) is considerably smaller 180 vs. 164. How should we interpret these figures? The first thing that comes to mind is that naming seems to be playing a more prominent role in English; while in Italian it is reference through a definite article in its deictic function that dominates C.'s production. As will be shown in chapter 8, the emergence of verbal predicates in Italian is surely a determining factor in the growth of deictic reference as opposed to simple naming.

In both languages, the fact that definites account for the largest number of articles overall is also due to the fact that, as Bennet-Kastor (1983:146) pointed out in her own study, the large number of definite articles on first mention of an NP "can be explained as generic or archetypal use, or as inclusive reference. This use increases with age". In other words the use of definite articles is justified on exophoric grounds. As the child's grammatical and communicative competence increases with age the referents of his or her universe of discourse become animated, they become subjects and objects rather than mere labels. If on the one hand this shows a growing mastery of the grammatical and pragmatic rules of discourse, on the other hand it also exposes one of the child's major weaknesses in the adult mastery of these rules: the ability to use definite articles anaphorically rather than deictically. In both languages C. mostly uses definite articles correctly, but only explores a subset of the possible uses of definite articles. Given the nature of the data collected, naturalistic and not experimental, it is unfortunate that there should be no independent control for C.'s actual knowledge of the anaphorical use of definites. In a way it is an artifact of the situation that he should be using so many deictic definite articles. After all shared knowledge can, and indeed must be inferred most of the time. In this light I believe that a controlled experimental situation with young bilingual children in the early phases of article acquisition would be more informative as to their actual competence. The data in C.'s corpus is however indicative of a parallel, if independent development of the semantic-pragmatic skills involved in article use, althogh other factors such as the earlier emergence of verbal predicates in a language and not the other surely contributes to the discrepancy of articles used in naming as opposed to other contexts.

6.9. Features and functional categories

In section 5.10, four predictions were formulated on the acquisitional sequence that children would follow in the acquisition of DPs. The analysis of C.'s use of definite and indefinite articles in Italian and English carried out in this chapter puts us in a position to provide results that will partially confirm the original predictions. The first prediction, repeated here for convenience stated that:

> (10)a initially, when children are not yet aware of the relevance of either agreement or referentiality features, they will produce a large proportion of determinerless nominals

Prediction (10)a is confirmed by our results in the distribution of articles in obligatory contexts over time. In both Italian and English there is an initial phase during which C. produces ungrammatical bare nouns. In four of the five earliest recordings in English all nouns appear ungrammatically without a determiner. In one of these first five English files, file 3 (1;11.4), C. produces 7 articles out of 44 obligatory contexts (15.09%), but it is only in file 7 (2;2.12) that for the first time a substantial proportion of nouns appear with an article, 13 articles are correctly supplied out of 21 obligatory contexts (61.9%) (see Table VII). In Italian it is only in the first two files that the proportion of articles in obligatory contexts is virtually equal to zero. From file 3 (1;11.17) until file 7 (2;0.23), correct suppliance of articles oscillates between 12.06% in file 6 (2;0.7) to 34.04% in file 7 (2;0.23). Although in both languages there is an initial period where bare nouns are the norm, this phase is more protracted in English than in Italian. Not surprisingly, given previous findings on the production of articles in children younger than two, prediction (10)a is confirmed.

Prediction (10)b concerns the initial stages of article production:

> (10)b the inital determinerless phase is followed by a stage in which children start to produce their first Art + N combinations; they are almost exlusively singular and lexically restricted to a small number of nouns

As already mentioned above, following an initial phase where all nouns are determinerless, in both English and Italian, C. starts to produce a number of Art + N combinations. As for the prediction that plural nominals will not be attested in the

child's early nominals, there is more evidence that this is the case in English than in Italian. In English there are no plural nominals until file 8 (2;2.24), while in Italian there is a modest number of plural Art + N combinations from the earliest files. In both languages there is a degree of lexical specificity whereby only a small number of nouns is initially found with an article. In English file 7 (2;2.12), out of 13 tokens of Art +N 6 are *a dog*, and 4 *a ball*. Until file 10 (2;4.29), only four nouns appear modified by an article in English (*dog, cat, ball, mouse*).

In Italian too, lexical specificity seems to play a role in C.'s first Art +N combinations. In file 3 (1;11.17), out of 19 Art + N strings 6 are accounted for by *l'orso, "the bear"*, 4 by *l'altro, "the other one"*, 3 by *l'altro poppode, "the other hyppo"*, and 3 by *il gatto, "the cat"*. A similar situation is found in file 4 (1;11.25) where *il lupo, "the wolf", l'altro orso, "the other bear", "il gatto"*, and *la scala, "the stairs"*, account for 10 of the 14 determinate nominals. As early as file 6 (2;0.7) however, C. proves to be able to extend the use of the article to a large number of familiar and new nouns. In file 6, for example, of the 13 noun types found in determinate nominals, only three appear more than once; and in file 7, out of 24 noun types in determinate nominals only five are repeated more than once. Although lexical specificity plays a role in both languages to kickstart C.'s article production, there is evidence that its effects last longer in English than in Italian.

Prediction (10)c concerns the marking of gender and number agreement in determinate nominals:

> (10)c gradually the number of nouns articles appear with increases, and children start to produce Art + N cominations more consistently. By this time there is still no clear evidence that articles are identified as markers of referentiality, but the local head-to-head relation between articles and nouns is appropriately marked by agreement.

Sections 6.3 and 6.5 have investigated the marking of gender and number agreement between articles and nouns in Italian, and number marking in English. The analysis of the Italian data shows that C. is remarkably accurate in the marking of both gender and number, the number of errors is extremely low overall (N = 15). Some of those errors are revealing overgeneralisations which provide interesting insights on the child's sensitivity to morphosyntactic cues in nominal agreement. A possible explanation for C.'s very good performance in marking these -Interpretable agreement

features was sought in the consistency and the transparency with which Italian marks gender and number on nouns and articles. In English too, although marking of plurality on nouns does not emerge untile file 8 (2;2.24), C. always correctly marks the plural feature on nouns, and uses the indefinite article only with singular count nouns.

At this stage, i.e. when production in obligatory contexts is 90% or over, gender and number are consistently marked as agreement features, and a large number of noun types are found in Art + N combinations, the child can be credited with an AGRN category. There is sufficient evidence with crediting the child with across-the-board notion of gender and number Agreement, and hence there is reason to believe that these formal -Interpretable features and their related AGRN projection are necessary to describe the child's competence.

As mentioned in section 5.1, the internal structure of DPs include a lower AGRN projection where non-interpretable phi features and Case are checked, and a higher D projection where the interpretable referentiality feature is checked. Moreover, we also illustrated how pragmatic competence is required in the choice of a definite vs. an indefinite article. The child must be able to abide by a discourse rule whereby previous linguistic discourse and knowledge of the interlocutor must be taken into account in the choice of article to use. The fourth of the predictions stated in 5.10 concerns this kind of semantic-pragmatic knowledge:

> (10)d the final stage is represented by the identification of the referentiality feature. Not only do children obligatorily satisfy the agreement relations between articles and nouns, they also use definite and indefinite aricles to identify referents appropriately

Our prediction was that the identification of the referentiality feature, and the development of the appropriate pragmatic knowledge would be the last thing to be acquired because of the cognitive demanding nature of this kind of knowledge. Suppose, like in C.'s case, that there is convincing evidence that the child has noticed lower-level morphological correspondences between articles and nouns with respect to gender and number agreement, and that he knows what a feminine, singular noun is, what a feminine, singular article is, and that these gender and number features on articles and nouns must agree with each other. This kind of knowledge is quite distinct from the kind of semantic and pragmatic knowledge that the child needs to use nominals referentially and appropriately according to the given/new pragmatic

distinction. In C.'s case, the data available cannot entirely confirm that the child has indeed reached this fourth stage where he has gained mastery of the above mentioned semantic and pragmatic skills. If this is the case, a D projection is missing in the child's representation of nominals, and its full realisation is subject to the identification of the function of referential DPs, and mastery of the discourse rule.

In conclusion, we have shown how C.'s acquisition of articles proceeds along two independent and separate routes. There is no transfer of features from one language to the other, and the earlier realisation of the obligatoriness of articles in Italian, together with the instantiation of gender and number agreement, does not seem to contribute to the acceleration of the same process in English, thus we have not found evidence of any bilingual bootstrapping mechanisms in this child's acquisition of nominals (Gawlitzek-Maiwlad & Tracy, 1996).

Chapter 7

English Verbs: formal features and lexical specificity

7.1. Introduction

In chapter 4 a number of hypotheses on the emergence of formal grammatical features, and their relation to the presence of Functional Categories (FCs) in early child grammar, were presented and discussed in some depth. The approach taken here favours a Weak Continuity approach whereby the presence of a FC is justified only by positive morphosyntactic evidence in the data. A general assumption of structural economy is at the heart of the WCH, according to which it is desirable to postulate only as much as structure as is necessary to account for the data at hand and no more (Safir, 1993; Grimshaw, 1994; Clahsen, 1996). In the structural analysis of child data, or for that matter of any linguistic data, it is licit to assume that a phrase structure position X^n is present only if X^n is phonetically realised or is involved in some syntactic process.

Of course, there may be cases in which the data is ambiguous, and more than one analysis can in principle be compatible with the empirical evidence. When this is the case the most economical analysis is to be preferred. Clahsen, Penke & Parodi (1994) propose the distinction between an analysis which is *compatible* with the empirical evidence, and one which is *supported* by it. If an analysis is compatible with the positive evidence, it accounts for all the data but assumes structure positions and/or features that are never phonetically realised and/or are never involved in any syntactic process. This type of analysis is consistent with the evidence, but makes structural assumptions that are not directly warranted by the data. By contrast, an analysis which is supported by the data is one which accounts for all the data in the most economical and minimal way, making the smallest set of assumptions that are needed to account for the positive empirical evidence. A Weak Continuity approach to language acquisition is

in a sense a more conservative data-driven approach that relies on direct positive evidence for any claim it makes about the structure of a child's grammar.

Another key element of a Weak Continuity approach is that FCs are not seen as monolithic entities, but as bundles of formal features, projections are head-driven by whatever features are realised in the grammar. The logical consequence of this view of syntax is that the focus in language acquisition should be primarily on the emergence of individual features, rather than on FCs *per se*. For example, in their study of longitudinal data from 7 monolingual German-speaking children, Clahsen, Penke & Parodi (1994) argue that the positive evidence that they review does not justify the assumption that the children's grammar at Stage I (MLU < 1.75) includes a C category, as claimed by Weissenborn (1990), Boser, Lust, Santelmann & Whitman (1992) and Poeppel & Wexler (1993). The fact that children distinguish between two categories of verbs, infinitival forms that appear in sentence-final position, and morphologically marked forms that tend to appear in V2 position, does not automatically imply that the V2 position is the Complementiser position.

In keeping with the principle of structural economy they advocate, Clahsen and colleagues prefer to call F this second position reserved for morphologically marked verbs. It is simply a functional position to the left of the VP where verbs that are beginning to be marked as finite move to. Because the Agreement paradigm has not been acquired yet, there is no reason to believe that the Agreement feature is relevant in the children's grammar at this stage, what seems to be spelled out is simply some notion of finiteness, the emerging feature is [± finite]. This being the case, the most minimal assumption that accounts for the emerging distinction between a finite class of verbs in V2, and a non-finite class of verbs in sentence-final position, is the postulation of a functional category where the [+ finite] feature can be checked. Because these children's verbs do not carry Agreement or Tense features at this initial stage, there is no need to assume that an AGR and/or a T category is present in the grammar. Moreover, because the C category in the adult grammar is also the category where [wh-features] are checked and no complementisers or wh-pronouns are present in the children's grammar at this stage, there is no reason to believe that [wh-features] are part of their grammar yet, and therefore a F category is all that is needed to check the only feature that these children's verbs actually spell out.

This line of argumentation along minimal and economical assumptions is the one that will be adopted in looking for evidence of formal features and FCs in C.'s data.

In addition a strict criterion of lexical productivity will also be adopted, whereby some degree of lexical flexibility in the production of inflectional morphology and in the choice of subjects and complements of a verb is required before crediting the child with the relevant morpho-syntactic knowledge.

The corpus was automatically searched for utterances containing the following verbal elements: copula *be,* auxiliary *be,* auxiliary *do*, auxiliary *have*, modals, and lexical verbs. Section 7.2 will deal with C.'s use of the copula , 7.3 with the aspectual auxiliary *be*, 7.4 with modals, 7.5 with the auxiliary *do*, 7.6 with lexical verbs and auxiliary *have*, 7.7 with a discussion of the data, and 7.8 with an analysis of C.'s subjects, 7.9 presents brief conclusions.

7.2. Copula *be*

Obligatory contexts for both auxiliares *be* and *have* and lexical *be* and *have* have been calculated by summing the number of actual occurrences and the number of contexts where they are obligatorily required in the adult language. For copula *be* the following contexts were searched: utterances containing an expletive subject pronoun (*it/there*), a deictic demonstrative pronoun (*that/those* or *this/these*), a personal pronoun (*I, you, s/he,* etc.), or a subject nominal (whether a proper name, or a common noun). The results for present tense copula *be* are summarized in Table I:

File	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.	% in o.c.
1	-	-	0/6	-	-	-	0
2	-	-	0/7	-	-	-	0
3	-	-	2/16	-	-	-	12.5
4	-	-	0/14	-	-	-	0
5	-	-	0/2	-	-	-	0
6	-	-	0/2	-	-	-	0
7	-	-	0/62	-	-	-	0
8	-	-	0/45	-	-	-	0
9	-	-	7/19	-	-	0/2	33.3
10	-	-	17/20	-	-	-	85
11	-	-	22/32	-	-	-	68.75
12	-	-	24/26	-	-	2/2	92
13	-	-	5/9	-	-	-	55.5
14	-	-	15/15	-	-	1/1	100
15	-	-	10/16	-	-	10/10	76.92
16	-	-	8/8	-	-	10/10	100
17	-	-	17/23	-	-	-	73.91
18	1/1	-	15/20	-	-	-	76.19
19	-	-	12/15	-	-	1/1	81.25

Table I: Present tense copula *be*

As shown in Table I, the only two contexts that have any significant representation in the data are 3 p.s. and 3 p.p., with the exception of 1 token of 1 p.s. in file 18. This is not surprising considering that 3 p.s. occurrences of the copula appear in one of two frames: *that's (a) x, it's (a) x,* and that 3 p.p. tokens are of the form *they're x.* All of these frames are involved in naming objects, and they are particularly frequent in the earlier stages when labelling is very prominent in C.'s speech. It is only in file 9 (2;4.7) that the copula makes its first appearance in the *that a x* construction. Previously C. simply used the deictic demonstrative pronoun *that* plus a noun optionally accompanied by a determiner:

(1) File 7
*CAR: that a dog .

(2) File 7
*CAR: tha(t) snake !

File 9 is also the file where two 3 p.p. contexts for the copula are found for the first time, but in both cases C. omits the copula, he also uses a singular demonstrative pronoun rather than a plural one:

(3) File 9
*CAR: dat *aminals .

(4) File 9
*CAR: dat two towers .

Although, overall 3 p.p. contexts only account for 6.78% of all utterances requiring an inflected form of the copula, performance in the plural is substantially better than in the singular, out of 26 obligatory contexts, the 3 p.p. copula is correctly supplied 24 times (92.3%). For 3 p.s. out of 357 obligatory contexts, the copula is only found 154 times correctly inflected (43.13%). The overall poorer performance in 3 p.s. singular contexts is however biased by an age effect, if one compares C.'s performance between 3 p.s and 3 p.p. contexts starting from file 9 when the first two 3 p.p. contexts appear, performance in the singular raises to 74.87%. Considering also that the obligatory contexts for 3 p.s. are almost 10 times the contexts for 3 p.p. (203 vs. 26), the figure of 74.87% correct suppliance in 3 p.s. obligatory contexts, reflects more

accurately C.'s growing awareness of the obligatoriness of the copula both in the singular and in the plural.

As mentioned above, 3 p.s. copula has a strong tendency to appear in one of two frames: *that's (a) x* and *it's (a) x*[1] There are no instances of the copula being used with a pronominal subject, and there is only one occasion where the copula is used with a proper noun:

(5) File 11
*CAR: <Marco is> [<] at school .

There is one example of a missing copula with a common noun:

(6) File 14
*CAR: that hand Alice .

In addition there are three ambiguous examples where it is not clear whether they are cases of a missing copula or, more likely, DPs where the adjective is postnominal, like in Italian, rather than prenominal, as required by English:

(7) File 8
*CAR: the mouse green .

(8) File 12
*CAR: the tunnel big!

(9) File 17
*CAR: yes it's a melon big .

Especially in example (9), it seems quite straightforward to analyse the utterance as a non-target N+Adj combination where the word order is Italian rather than English. The rest of the missing copulas are in *that's (a) x, it's (a) x* frames. Because the virtual totality of instances of 3 p.s. copula also occurr in *that 's a x, it's a x* frames, it would thus be reasonable to suggest that the copula in *that's* and *it's* is

[1]There re also 13 occurrences of *what's that?* (Files 10-18) and 4 occurrences of *where is it?* (Files 14-17) which can be simply considered as instances of rote-learned frozen forms. There are no instances in the data in which a subject other than *that* is found in the object copula *what*-questions, or other than *it* in the adjunct *where*-questions.

simply an unanalysed, frozen form which the child has learnt as part and parcel of the *that's* and *it's* chunks. There are indications that this analysis is not on the right track, if *that's* and *it's* were simply learnt as whole units, one would expect C. to use them not only in these frames, but also in other contexts: e.g. *I want that's dog* instead of *I want that dog*. These examples are unattested, on the contrary there is evidence that C. can use *that* and *it* as distinct from the *that's* and *it's* forms in a variety of occasions. Both *it's* and *it* , appear as subjects, and only *it* is ever found in direct object position:

(10) File 14
*CAR: it's not dark.

(11) File 16
*CAR: it's a person.

(12) File 17
*CAR: it goes.

(13) File 17
*CAR: I see it!

The same holds for *that* , *that's* is only ever found in subject position, while *that* is found both in subject and object position:

(14) File 16
*CAR: like that's the light.

(15) File 18
*CAR: no # that's a cow too.

(16) File 16
*CAR: Marco give that too.

(17) File 17
*CAR: that go there.

(18) File 19
*CAR: who gave me that?

The existence of both forms, *that's/it's* being used in copula constructions, and *that/it* being used both in copula constructions and also as subjects and objects of auxiliaries and lexical verbs, indicates that C. appreciates that *that's/it's* are a combination of two distinct elements, he does not use them in free variation with *that* and *it*. Moreover *that's* and *it's* make their appearance after a protracted initial period where any form of copula is totally absent. If they were truly unanalysed forms one might expect them to emerge at earlier stages as part of a frozen form. In fact it would not be surprising to find them from the very beginning of acquisition, although there is evidence from studies of phonological acquisition that the /ts/ cluster may well pose articulatory problems to the young child, and as such is unattested in the earliest phases.

There is additional evidence from occurrences of 3 p.p. copula that *they're* is used in the appropriate contexts and that C. uses *they* and not an unanalysed form *they're* as the subject of modals and lexical verbs:

(19) File 19
*CAR: they go faster to the door.

(20) File 18
*CAR: they can't go here .

If, on the one hand, there seems to be some evidence that C. can use the contractible copula, both 3 p.s. and 3 p.p., in the appropriate contexts, nevertheless there are no obvious indications that he can use 3 p.s. copula with subjects other than *that* and *it*, and 3 p.p. copula with a subject other than *they*. For the singular there is only one instance of the copula being used with a proper noun subject, and no instances of it with personal pronouns. For the 3 p.p. the only subject that ever appears with the contracted form *'re* is *they,* again no proper, common nouns or demonstrative pronouns are ever used with it. Although there is a gradual increase in the use of the copula from file 9 onwards, C. only uses three different subject types with it: *that, it,* and *they*.

It seems safe to conclude that *that's/it's (a) x* and *they're (a) x* function as productive semiformulaic frames where all of the elements have been analysed by the child as single words in their own right and are therefore not frozen forms. However,

although the single elements (i.e. the demonstrative pronoun *that,* the expletive pronoun *it,* the personal pronoun *they,* the optional determiner, and the following noun/adjective) also appear in other contexts, whenever the copula is used it is only used with these three different subject types. If one thinks of the frame as segmented into two slots: *that's/they're* and *(a) x,* it is clear that the degree of lexical specificity is by far higher in the first slot where only three different types of subjects are allowed; the second slot can host any determiner and noun combination. C.'s copula constructions are therefore a prime example of what Lieven, Pine & Baldwin (1997) call lexically-based patterns, productive positional patterns in which a constant item, in this case *that's/it's/they're,* occupies a constant position, here a construction-initial position, with respect to variable items with which it appears, in this case the determiner + noun combinations.

To be more accurate, it is only the plural copula constructions that entirely fit the description of a lexically-based pattern, in the case of the singular copula constructions there is a minimal degree of flexibility in the type of lexical material that can appear in the fixed construction-initial position: either *that's* or *it's* . Of course one might argue that C. has two different lexically-based patterns, one with *that's* + determiner + noun and one with *it's* + determiner + noun. However, because of the overlap of the contexts in which the two patterns occurr, it is reasonable to assume that it is the same positional frame where the constant initial position is not absolutely specified for only one lexical item, but shows a minimal degree of flexibility in allowing two different lexical items to appear there. *That* and *it* are used virtually interchangeably as subjects of copula constructions, but it is interesting to note that all instances of missing copula, both in the first 8 files when the copula has not emerged at all yet, and in the subsequent files when the copula is gradually more and more frequent in obligatory contexts, are of the type *that (a) x.* Whenever *it* is in subject position, the copula is always obligatorily supplied, and there are no occurrences of *it (a) x.*

The nature of the lexical specificity of C.'s copula constructions is to be explained by the kind of discourse function that such constructions serve: in essence C. only ever uses the copula for labelling, for naming objects. From the earliest files, where the majority of his verbless utterances are of the type *that (a) x,* his main mode of verbal interaction is through naming objects around him, initially by using an all-purpose deictic such as the demonstrative pronoun *that* plus a determinerless noun: *that x.* Subsequently, determiners start to appear in the frame: *that a x,* and at a later

stage (File 9 onwards) the copula starts being produced: *that's a x.* At the same time as the copula makes its first appearance, *it's* begins to emerge and to be used in free variation with *that's*. Tracking the development of labelling contexts over time we observe the original positional pattern deictic + noun gradually evolving into the adult copula construction where both halves are now target-like: the deictic is combined with the copula and the noun is preceded by an article.

Also noteworthy is the fact that C. never makes any Agreement errors when the copula is supplied, a 3 p.s. subject is always correctly used with a 3 p.s. copula, and similarly a 3 p.p. subject is always followed by a 3 p.p. copula. All of C.'s errors with the copula are omission errors, the absence of commission errors indicate that he must be sensitive to the distributional properties of the inflected forms of *be*: *is* in 3 p.s. contexts and *are* in 3 p.p. contexts.

In the data there are also 10 questions with an inflected form of copula *be*, two are subjectless and it is therefore impossible to assess whether any inversion has taken place:

(21) File 16
*CAR: where [/] where is ?

(22) File 19
*CAR: and what is called ?

In the remaining 8 questions, subject-auxiliary inversion is correct in 5 cases, the remaining ungrammatical examples are given below:

(23) File 12
*CAR: this is chocolate?

(24) File 15
*CAR: where they are?

(25) File 15
*CAR: where they are?

While example (23) could be marginally grammatical as an echo question, although from the context it is unlikely that C. intended it to be an echo question. The other two examples where the copula has failed to raise to C to check its wh-feature are decidedly ungrammatical. Because of the very limited amount of data on questions containing a copula and a subject, it is difficult to draw any definitive conclusions about the status of the wh-feature carried by the verb in these questions. What is clearer is that a wh-operator is always in sentence-initial position, thus satisfying the requirement that a wh-operator be in a scope position (Rizzi, 1991).

The few examples of negative copula constructions (4) are all of the type *it's + not:*

(26) File 14
*CAR: it's not dark .

(27) File 17
*CAR: no # it's not grandma mouse .

Although the negative marker is always correctly placed after the copula, because there are only examples of contracted *'s,* and their number is very limited it is not possible to draw any definitive conclusions about C.'s knowledge of the hierarchical position of inflected verbs and negation.

The evidence reviewed above supports the claim that C. can use inflected forms of the copula *be* in an appropriate fashion. However, the fact that he only uses it in a very limited range of lexically specific syntactic contexts, combined with his very limited knowledge of the morphological paradigm, raises some doubts as to C.'s actual mastery of the copula and its associated Agreement and Tense features.

Similar conclusions on the analysis of the copula in ECG are reached by Radford (1990) in an in-depth analysis of a large corpus of utterances from monolingual English children. Radford puts forward five main reasons why early instances of copula *be* cannot be taken as evidence that English-speaking children do in fact treat the copula as a sort of dummy auxiliary spelling out Agreement and Tense features. Firstly, at the same time as children produce tokens of clitic copula *'s*, they do not display any productive control of singular *-s* inflection on lexical verbs. If use of the copula really meant that the child appreciates what it stands for, one should expect

that this realisation should hold across the board and that Agreement and Tense features should be made visible through appropriate inflectional morphology also on lexical verbs and on other auxiliaries. This is not the case, neither for Radford's corpus, nor for C.'s corpus, no productive use of inflectional morphology is attested at this stage.

Secondly, 3 p.s. copula forms account for the overwhelming majority of copula contexts; there is no evidence that other forms of the morphological paradigm are used at any significant level, virtually no instances of *am/are/was/were*. This is equally true for C.'s data, with the exception of a number of clitic 3 p.p. forms attested from file 12 onwards (2;9.6), and 1 1 p.s. form in file 18, the rest are all 3 p.s. forms. Moreover the 3 p.p. copula forms appear approximately 5 months later than the first 3 p.s. forms are attested (2;4.7).

Thirdly, the percentage of *'s* forms in obligatory contexts in C.'s data reaches 100% only in two files (file 14, and file 16), in the rest of the corpus, from when the copula is attested with some consistency (file 9) to file 19, the average is 71.42%, by the end of the data collection period there is still some degree of optionality, omissions still occur an average of 30% of the time. Unfortunately Radford does not give quantitative data of copula suppletion in his corpus, but he states that "the *s*-forms are used only sporadically, alongside *s*-less forms" (Radford, 1990: 166). C.'s use is clearly more than "sporadic", but it is not consistently target-like.

A fourth reason that Radford gives for not treating the contractible copula as a sign of adult-like mastery of Agreement and Tense features is the fact that forms such as *there's/where's* are used not only with singular complements, but also incorrectly with plural complements. There is no such evidence in C.'s data, the only errors found are omission errors, there are no commission errors. A fifth and final argument is that the range of lexical item to which the copula is suffixed is typically limited to a small set of inanimate pronouns (*it, that, what*) and to the locative pronouns *there* and *where*. As shown by the detailed analysis of the data above, this is also the case for C., only *that* and *it* ever appear as subjects to *'s*, there is only one case in which the subject is a proper noun (example (5)), and no cases of personal pronouns or lexical DPs. As for the cases in which the copula is 3 p.p. only one type of subject is attested: the personal pronoun *they*.

Radford's observations on monolingual English children's use of the copula are clearly relevant to C.'s data as well. Given the obvious limitations that apply to the child's data, the conclusion drawn here are parallel to Radford's: there is no clear indication that the copula is anything more than an element in a productive positional pattern. The evidence reviewed here and in Radford (1990) leads us to conclude that the copula in these initial stages does not spell out Agreement and Tense features.

7.3. Auxiliary *be*

Data on C.s performance with the copula indicate that his knowledge is confined to a very small set of contexts of use, only 3 p. contexts are represented in the data. The copula *be* has a homonym in the progressive auxiliary *be*, but even from a cursory look at the data in Table II, it is clear that the child treats the auxiliary very differently from the copula. There are indications that the variety of contexts in which the auxiliary appears, or at least the contexts in which it is required, are somewhat greater. Table II reports data for the aspectual auxilary *be* in present tense contexts:

File	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.	% in o.c.
1	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-
8	-	-	0/2	-	-	-	0
9	-	-	-	-	-	-	-
10	-	-	-	-	-	-	-
11	-	-	3/3	-	-	-	100
12	-	-	19/22	-	-	-	86.36
13	0/1	-	-	-	-	-	0
14	0/1	-	18/21	-	-	2/2	83.33
15	0/1	-	17/17	-	-	2/2	95
16	-	1/1	3/3	1/4	-	-	62.5
17	-	-	5/6	-	-	2/3	77.77
18	-	0/2	4/4	1/1	-	-	71.42
19	-	-	2/2	-	-	2/2	100

Table II. Present tense auxiliary *be*

Auxiliary *be* contexts start appearing in File 8 (2;2.24) where there are two subjectless and auxiliaryless questions where the subject and the auxiliary should have been 3 p.s.:

(28) File 8
*CAR: what doing ?

(29) File 8
*CAR: what doing ?

It is only in File 11 (2;7.8) that the first three tokens of 3 p.s. *be* are to be found:

(30) File 11
*CAR: <what's he doing> [<] ?

(31) File 11
*CAR: <what's he doing> ?

(32) File 11
*CAR: is hiding .

The presence of the auxiliary in the two object questions above is probably part of a semiformulaic utterance. Although there is evidence that from File 12 (2;9.6) onwards there is some productive use of 3 p.s. auxiliary *be*, there are 11 tokens of *what's x + -ing* form*?* questions throughout the corpus, and the only verb that appears with the progressive *-ing* suffix in these questions is *doing*. In declarative sentences, C. uses the auxiliary *be* with 15 different types of verbs in the progressive form thus showing a certain degree of flexibility (*hiding, doing, crossing, going, coming, bringing, falling, crying, swimming, getting, putting, making, playing, eating, raining*).

By contrast with the copula *be*, where the only contexts are 3 p.s. and 3 p.p., 5 out of 6 possible contexts are represented in the data for aspectual auxiliary *be*. Not surprisingly the only missing context is that for 2 p.p., given that most of the time C. is alone with only one adult, the number of occasions in which he could use a 2 p.p. to address two or more people present at the same time are drastically reduced. The fact that 5 contexts are found in the data does not automatically mean that overt forms of the auxiliary are supplied in all of these five contexts. For instance there are no attested occurrences of 1 p.s. Whenever it is required C. fails to supply it:

(33) File 13

*CAR: I going to put it there .

(34) File 14
*CAR: I going to have a butterfly .

(35) File 15
*CAR: I making a stairs.

Only one out of three required instances of 2 p.s. is correctly provided:

(36) File 16
*CAR: you're making a *three!

And out of five contexts for 1 p.p. only two are attested:

(37) File 16
*CAR: we're making a *thrain.

(38) File 18
*CAR: we're making a house.

Performance is instead virtually perfect for 3 p.p. contexts, 8 out of 9 forms are correctly produced (88.88%). Equally good is C.'s use of the auxiliary *be* in 3 p.s. contexts: 71/80 (88.75%). Although contexts other than 3 p. are represented in the data, and correct forms of the auxiliary *be* are provided 3 times out of 11 obligatory contexts, the vast majority of occurrences are still 3 p.: 79 out 82 forms are either 3 p.s. or 3 p.p.. When looking in more detail at the function that utterances containing instances of 3 p. auxiliary *be* perform, one finds that they can be divided into two groups: one group includes questions of the type *What's x doing?* (12 tokens), the rest are descriptions of on-going activities involving one or more participants:[2]

(39) File 14

[2] In File 14 there is also one example of a why question:

File 14
*CAR: why he is falling?

*CAR: what's she doing ?
%exp: asking Eric about a picture of Alice in the book

(40) File 14
*CAR: and what's he doing there ?
%exp: asking Eric about a picture of Alice in the book

(41) File 15
*CAR: that is going <to the> [/] to the tunnel .
%act: pushing the toy train towards Eric's knees.

(42) File 17
*CAR: they're going to the dad
%act: moving two toy mice towards a third bigger mouse.

(43) File 18
*CAR: <he's eating> [/] he's eating all the fish
%act: Carlo makes the toy cat eat the fish.

For 3 p.s. there is evidence that C. knows both the contracted form and the uncontracted form, he uses both, although there is a higher proportion of contracted forms (52 contracted forms vs. 19 uncontracted forms). For 3 p.p. there are only instances of the contracted form cliticised onto the pronoun *they* in *they're*. One might therefore suspect that these may simply be unanalysed forms, and not true instances of a pronoun plus a cliticised form of the auxiliary. There is however evidence that C. knows the form of the pronoun *they* and that he can use it in appropriate contexts as the subject of modals and verbs:

(44) File 15
*CAR: I know where they are .

(45) File 18
*CAR: they can't fit .

The same pattern holds for 2 p.s. and 1 p.p.. Athough the 3 forms with auxiliary *be* in the corpus are instances of contraction, there are numerous examples of both *you* and *we* being used correctly as subject pronouns:

(46) File 14
*CAR: we finished .

(47) File 18
*CAR: we lost the penguins .

(48) File 15
*CAR: you have to put this .

(49) File 17
*CAR: you want to help me ?

Because C. can and does use personal pronouns without a cliticised auxiliary *be* when these serve as subjects of other verbs, there is no reason to treat the contracted forms of the auxiliary as completely unanalysed forms. Although there is a very limited number of them, only 11 in the whole corpus, nevertheless it is reasonable to assume that the child is aware that the *'re* is a separate element, and that as such *we* and *we're* are very distinct forms.

Only two examples of a negated auxilary be are found throughout the corpus:

(50) File 19

*ERI: is it gonna make it ?
*CAR: it's not !

(51) File 14
*CAR: is not crying .

As for the copula *be*, there are no commission errors, only omission errors. Although the 1 p.s. auxiliary *be* is never supplied in the three contexts where it is required, C. simply omits it, he never uses a wrong form as for example *@I is* or *@I are*. The same goes for 2 p.s. and 1 p.p., whenever the auxiliary is present it is in its correct form *are*, there are no cases of *@you is* or *@we is*.[3] Whenever a form of the

[3]Ingham (1998) reports a high number of commission errors in a longitudinal study of a monolingual English child, Sophie (2;5-3;0). Not only does Sophie commit omission errors with 3 p.s. lexical verb forms where she fails to supply the *-s* inflection, she also does not seem to treat suppletive forms like *is* and *are* as agreeing forms. Sophie uses 3 p.s. and 3 p.p. forms of copula *be* and of

auxiliary *be* is present it is correctly inflected for person and number, this indicates some awareness on C.'s part of the requirement that the nominal phi features of the subject must be checked by an auxiliary carrying those same phi features. It must however be pointed out that C.'s knowledge of the morphological paradigm is somewhat limited, and that out of 82 inflected forms of the auxiliary *be*, 12 (14.63%) are combined with the verb *doing* (11 of which are found in the semiformulaic question frame *what's X doing?*), and 40 (48.78%) with the verb *going*. These two verbs together account for 63.41% of all progressive forms combined with the inflected auxiliary *be*. It is a fact that C. uses these progressive forms with the auxiliary with correct subject-verb Agreement, however the high degree of lexical specificity in these progressive constructions is by no means a negligiable factor.

The next two sections assess C.'s use of modals and of auxiliary *do* in negatives and interrogatives. Once more, the degree of lexical specificity in the child's use of these verb forms will be investigated to evaluate to what extent his performance can be said to be the productive result of adult-like competence.

7.4. Modals

This section reports on C.'s use of modals and auxiliary *do*. In Table III all tokens of modal verbs are collapsed. Because it is virtually impossible to identify obligatory contexts for modals, only the number of occurrences were counted. Four different verb types are found: *can, could, might,* and *should.*

---

auxiliary *be* indifferently for singular or plural contexts, there are no Number distinctions, and equally in non-3 p. contexts, no Person distinctions.

File	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	-	-	-	-	-	-
4	-	-	-	-	-	-
5	-	-	-	-	-	-
6	-	-	-	-	-	-
7	-	-	-	-	-	-
8	-	-	-	-	-	-
9	-	-	-	-	-	-
10	-	-	-	-	-	-
11	-	-	-	-	-	-
12	-	-	-	-	-	-
13	22	-	-	-	-	-
14	-	-	1	-	-	-
15	4	-	3	-	-	-
16	-	-	-	1	-	-
17	1	-	-	-	-	-
18	-	-	9	7	-	8
19	2	1	13	-	-	-

Table III. Number of modals

As can be seen from the table above modals make their first appearance in file 13 (2;10.1), are used more sporadically from file 14 (2;10.15) to file 17 (3;0.3), and are produced again with some consistency in the last two files. Except for two occurrences of *couldn't* in file 19 (3;1.25) as the past of *can*, all other instances of modals are used in the present tense:

(52) File 19
*CAR: because # <I couldn't> [/] I couldn't help Rachel.

In file 13 (2;10.1), there are 22 instances of the verb *can* with a pronominal subject followed by the verb *put:*

(53) File 13
*CAR: I can put this flower .

The only modal appearing in file 14 (2;10.15) is a negated form of *can* preceded by a 3 p.s. pronominal subject:

(54) File 14
*CAR: he can't go into the door .

In file 15 (2;10.23) there are again two examples of *I can put (a) x,* one example of *I can* followed by the verb *found* (a wrong lexical choice where *found* stands for *find*), and one instance of *I can't* followed by *found.* The three 3 p.s. forms are of the type *pronominal/nominal subject can't go:*

(55) File 15
*CAR: it can't go!

(56) File 15
*CAR: but the bear can't go!

In file 16 (2;10.30) we find the first example of inverted modal auxiliary in a yes/no question:

(57) File 16
*CAR: can we open the door?

Only one occurrence of modal *can* is attested in file 17:

(58) File 17
*CAR: I can see a bed .

A variety of modals are found in the last two files. In file 18 (3;0.16) there are 3 instances of *can* with a 3 p.s. subject followed by the verb go:

(59) File 18
*CAR: this can go up and down .

The remaining 6 occurrences of a 3 p.s. modal are represented by the negated modal *can't* followed by *go* (1 time) and *fit* (5 times):

(60) File 18
*CAR: <they [/] they [//] this can't fit there> [<] .

Of the 7 occurrences of a 1 p.p. modal, 6 are instances of negation, while 1 is a positive declarative sentence:

(61) File 18
*CAR: we [/] we can't find them .

(62) File 18
*CAR: we can put this

The 7 examples of 3 p.p. are tokens of *can*, 4 of which are in the negated form followed by *go* (4 times) or *fit* (once). The 3 positive declarative utterances containing a modal are all of the type *subject + can go* :

(63) File 18
*CAR: all the giraffes can go .

In file 19 there are the only two examples of a modal in the past tense (see example (52) above), and the only occurrence of a modal with a 2 p.s.:

(64) File 19
*CAR: can you help me ?

For the first time modals other than *can* also appear in file 19, out of the 12 3 p.s. there are 3 tokens of *can,* 4 examples of can't, 1 instance of *should,* 1 example of *shouldn't*, 1 *could* (present conditional), and 2 *might.*

Table IV summarises the distribution of modals and the verbs they take as complements:

Person/Number	Modal verb	Verb types	Verb tokens
1 p.s.	can	put	24
		found (= find)	2
		see	1
	couldn't	-	1
		help	1
2 p.s.	can	help	1
3 p.s.	can	go	3
		see	2
		be	1
3 p.s.	can't	go	5
		fit	5
		fall	2
		be	1
		see	1
3 p.s.	might	fall	1
		go	1
3 p.s.	could	be	2
3 p.s.	should	be	1
3 p.s.	shouldn't	fall	1
1 p.p.	can	put	1
1 p.p.	can't	-	3
		look	3
		find	1
3 p.p.	can	go	3
3 p.p.	can't	go	4
		fit	1

Table IV. Distribution of modals and their verbal complements

As shown in Table IV, one single modal, *can,* appears in 88.88% of all utterances containing a modal verb (64/72). Moreover, two verbs alone, *put* (25 tokens) and *go* (16 tokens), account for more than 50% of all modal verbal complements. Furthermore, if one looks at the distribution of verbal complements that *can* takes, it can be easily seen that the degree of lexical specificity with which lexical verbal complements appear is even higher than for all the modals collapsed. All of the 25 tokens of *put* and the 16 tokens of *go* appear as the complements of *can*, out of a total number of 57 different occurrences of this modal. In other words, there is a 71.92% chance that whenever *can* is attested in the corpus it will take either *put* (43.85%) or *go* (28.07%) as its complement. Given that, at least until file 19 where other modal types are attested (*should, shouldn't, could,* and *might*), *can* is the only representative of what is typically defined as the modal category for adult English, and considering also that this verb appears with an extremely limited range of complements, it seems premature to speak of the existence of a modal category in C.'s grammar at this stage. However, although from a purely distributional point of view, there seems to be little positive evidence from C.'s limited use of only one modal verb with a very lexically-specific range of complements that a modal category is established, there is some

indication that can and the other few examples of modals are treated as a separate class from lexical verbs.

From a syntactic point of view, in the 4 occasions where a modal appears in a yes/no question it is correctly inverted 3 times. As far as placement of modals with respect to negation is concerned, it is more difficult to draw any definitive conclusions. Whenever a modal is supplied in a negative sentence, the negation is always cliticised onto it there are no instances of *can not* or *should not*, only *can't* and *shouldn't*. It could reasonably be argued that the modals with cliticised negation could have been learnt by C. as whole lexical items which are only used in contexts where he wants to express the negation of a modal, but this does not necessarily prove that he knows the syntactic distribution of modals to the left of negation. However, it is also true that there are no instances of negative sentences with a negative marker cliticised onto a lexical verb. In the few examples of negation other than the frozen form *I don't know* C. uses a *do* form with a cliticised negative marker before the lexical verb:

(65) File 16
*CAR: I want to break [?] I don't want to make a tunnel.

There are 4 questions containing the modal *can*, and in three of them the modal is correctly inverted:

(66) File 13
*CAR: I can put?

(67) File 13
*CAR: where Kain@u [= can I] put it?

(68) FIle 16
*CAR: can we open the door ?

(69) File 19

*CAR: can you help me ?

Moreover, verbal complements of modals are always bare verbs, they are never introduced by the infinitival marker *to*, which is correctly used in the complement clauses of verbs other than modals:

(70) File 15
*CAR: I want to make a garden .

(71) File 15
*CAR: you have to move !

(72) File 18
*CAR: I like to build a house .

(73) File 19
*CAR: no # we've got to put it here .

The negation facts, together with correct inversion in 3 out of 4 questions, and the selection of complement clauses with a bare verb, suggest that C. does in fact treat *can* and a handful of other verbs (*should/shouldn't, could,* and *might)* as a separate class of verbs from lexical verbs. Although C. does not seem to have generalized the syntactic properties of *can* to a whole class of verbs, he is accurate in the use of this one modal verb, his rather conservative lexically-specific strategy is such that he is more likely to commit an error of omission rather than one of commission.

Another interesting property of modals in C.'s corpus is that they always have an overt subject in the Nominative Case, be it a pronoun or a lexical DP. There are no instances in which a modal verb is produced with a null subject. On the assumption that English modals must check their features in T, and assuming that Nominative Case must be checked in spec-AGRS, it could be argued that C.'s modals and their Nominative subjects can be taken as evidence for the presence of both T and AGR. Although this is a possibility, nevertheless the extremely high degree of lexical specificity in the distribution of subject, modal verb, and verbal complement is an indication that C.'s knowledge about modals and their properties is still tied to a very small number of specific lexical items. The following section on the use and distribution of auxiliary *do* will further weaken the argument in favour of FCs in C.'s grammar at this stage.

7.5. Auxiliary *do*

Auxiliary *do* carries Agreement and Tense features that cannot be overtly checked by the raising of lexical verbs which have weak Agreement and Tense features in Modern English. In negative sentences the presence of an intermediate Negative head between VP and TP/AGRP blocks percolation of the weak Agreement and Tense features on the lexical verb in V. The auxiliary *do* is therefore inserted to the left of Negation so that the relevant -Interpretable Case and and phi-features can be checked and erased with the subject in spec-AGRP.

In questions, a strong +wh feature generated on T must be overtly moved to C to be checked in a spec-head configuration with a wh-operator.[4] Because English lexical verbs cannot move to T, and hence cannot move to C, an auxiliary *do* is inserted in the derivation so that the strong wh-feature can be checked by overt movement of the auxiliary to C. Because of the last resort nature of *do*, it is simply a way that formal grammatical Tense, Agreement, and wh-features features can be spelled out and appropriately checked, so that the derivation can converge.

The presence of inflected forms of *do* in questions and negative sentences is an indication that the child is aware that these Agreement , Tense, and wh-features exist and must be appropriately checked by the insertion of a dummy element *do* . Because the child is assumed to have innate knowledge of the Principle of Full Interpretation, as part of UG, and in the absence of any evidence that English-speaking children ever entertain the possibility that English lexical verbs have strong features, if the child fails to supply *do* where required, it seems reasonable to assume that for the child at this stage verbs do not carry the relevant Agreement and Tense features that they have in the adult grammar. Because FCs are instantiated on the basis of morpholexical evidence, it follows that, at a stage when the child has not yet discovered that verbs carry Agreement and Tense, the child's grammar does not include AGR and T.

Data on C.'s use of *do* in interrogative and negative contexts, and lack thereof, shows that there is little evidence substantiating the existence of AGR and T categories. Table V reports on the use of *do* in matrix non-subject questions:

[4] See Rizzi (1991, 1997) for the formulation of the Wh-Criterion.

File	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.	% in o.c.
1	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-
7	-	-	0/1	-	-		0
8	-	-	-	-	-	-	-
9	-	0/1	-	-	-	-	0
10	-	-	-	-	-	-	-
11	-	-	-	-	-	-	-
12	-	-	-	-	-	-	-
13	-	-	-	-	-	-	-
14	0/3	-	-	-	-	-	0
15	-	-	-	-	-	-	-
16	0/2	-	-	-	-	-	0
17	-	0/2	0/1	-	-	-	0
18	-	-	1/2	-	-	-	50
19	-	-	-	-	-	-	-

Table V. *Do* in matrix non-subject questions

There is only one occurrence of auxiliary do in C.'s main questions with a lexical verb:

(74) File 18
*CAR: why do not fit?

Out of a total of 12 questions there are 2 yes/no questions, 2 *what* object questions, 5 *where* adjunct questions, and 1 *why* question. With the exception of a question in file 17 where not only the auxiliary, but also the subject is missing, all the other questions have an overt preverbal subject with the exception of a question in file 7, where the subject is postverbal:[5]

(75) File 7
*CAR: what say Bambi ?

The subjectless/auxiliarless yes/no question in file 17 has a recoverable 3 p.s. subject:

[5] This postverbal subject is in all likelihood an interference from Italian where postverbal subjects are allowed both in declarative and interrogative sentences.

(76) File 17
*CAR: go there ?

In addition to the two 3 p.s. questions in (70) and (71) above, there is an adjunct 3 p.s. question in file 18:

(77) File 18
*CAR: where this donkey go ?

In all of three cases the lexical verbs *say* and *go* are uninflected, they are simple bare forms. This type of auxiliaryless question contrasts with English child questions that have been reported in the literature where there is no auxiliary, but the main verb is inflected for Agreement and/or Tense (cf. Stromwold, 1990; Guasti, 1996; Guasti & Rizzi, 1996):

(78) What John eats/ate?

C. does not use inflected forms in any of his three questions with 3 p.s. subject, there is reason to assume that these, together with the main verbs found in the other auxiliaryless questions are bare non-finite forms. Guasti (1996) and Guasti & Rizzi (1996) have proposed that in auxiliaryless questions of the type in (70), there is a null counterpart of *do*, whose only difference from lexical *do* is that it is phonologically null. In essence a question such as *Where this donkey go?* is adult-like in every respect, except for the fact that the child allows a null auxiliary in C as opposed to a lexical one. The reason why this option is licit in child grammar, but not in the adult grammar, is that children can omit the topmost layer of the clause, what Rizzi (1997) calls the ForceP. In doing so their clause only projects as far as FocP (what is normally called CP), if ForceP is not present above FocP, the auxiliary can be null because it is in the head of the root and in this position it is not subject to a principle of syntactic identification. Empty categories need to be identified through chain-connection to an antecedent if and only if such an antecedent exists. If the null element in question is in the root of the clause it is in a position where it cannot be c-commanded by any antecedent, and therefore no syntactic chain need to be made visible (cf. Rizzi, 1994a, 1994b for a similar proposal for null subjects and root infinitives respectively).

Alternatively, one may want to argue that at the stage at which C. produces auxiliaryless questions such as *Where this donkey go?*, there is no need to assume that

an auxiliary, be it overt or null, is necessary to carry Agreement and Tense features that are not yet part of the child's grammar at this stage. By a process of Merge C. is simply combining a main clause containing an uninflected bare verb with a wh-word. Roeper (1996) makes just such a proposal for non-inverted wh-expressions (from McNeill, 1970):

(79)a. where the other Joe will drive?
b. what he can ride in?
c. what he can do?

Roeper (1996: 427) states that the examples above "could be a straightforward case of Merger between a wh-word and an IP". It is only when the child realises, on the basis of evidence from uninverted embedded questions, that the wh-word is in spec-CP that the need for a C projection is finally realised. Roeper also points out that this reanalysis is not an abrupt across-the-board phenomenon, but rather a process that takes place on an item by item basis. He proposes a two-stage developmental process:

(80) Stage I:
a. NP object: *ask John*
b. No inversion: *what he can sing*
Stage II
a. Indirect Q complement: *ask John what he wants*
b. Inversion: *what does he want?*

In the Adam corpus (Brown, 1973), Roeper notes that the SAI in non-subject questions occurs around file 15 for *how* questions, around file 20 for *what* questions, and around file 55 for *why* questions. In C.'s data there is no indication from questions involving a bare form of a lexical verb that Agreement, Tense, and/or wh-features are represented in the child's grammar in an adult-like fashion by the end of the recording sessions (age 3;1.25).

As already seen in section 7.3, there is also little evidence that inflected form of the auxiliary *be* in questions are truly examples of raising to C. Although an inflected 3 p.s. form of the auxiliary *be* is correctly produced in 14 out of 19 obligatory contexts, in two cases the question is subjectless and therefore SAI is impossible to test:

(81) File 14
*CAR: what is doing ?

(82) File 14
*CAR: is swimming ?

Of the remaining 12 cases, 11 are embedded in the frame *what's he/she doing?,* and one is a *why* question where the copula has failed to raise:

(83) File 14
*CAR: why he is falling ?

As far as negation is concerned, once again there is no strong evidence that C. has analysed the auxiliary *do* and the clitic negative marker *'nt* as two distinct elements. What is more, negation only ever appears with the verb know in the frozen fom *I don't know*. Throughtout the corpus there are 21 examples of *I don't know,* the first occurrence being registered in file 10. There are also two examples with *I don't know* + a complement:

(84) File 14
*CAR: I don't know that thing .

(85) File 18
*CAR: I don't know what is it .[6]

The only other verb that is found with *don't* is *want*::

(86) File 12
*CAR: I [/] I don't want this .

(87) File 16
*CAR: I want to break [?], I don't want to make a tunnel .

[6]This example of postverbal subject, this time in an indirect interrogative, is again likely to be the result of crosslinguistic interference from Italian. Such use of postverbal subjects is not attested in monolingual English children to my knowledge.

There is one isolated example of the negative marker *not* not being cliticised onto the auxiliary *do,* and it is in a subjectless negative question where the missing subject has a 3 p.s. referent, and therefore *does* and not *do* is required:

(88) File 18

*ERI: that should fit .
*CAR: that can't fit .
*CAR: xxx there .
*ERI: yeah ?
*CAR: no .
*ERI: no .
*CAR: oy # why [/] why do not fit ?

Doesn't is attested on one occasion in file 16 in a declarative with a missing 3 p.s. subject:

(89) File 16
*CAR: no, doesn't go there .

In file 12 there are two negative imperatives where *not* and *don't* are used interchangeably:

(90) File 12
*CAR: oh no # don't put that there .

(91) File 12
*CAR: no # no put that there okay ?

Finally in file 16 don't is incorrectly used with *got,* where *got* may be misanalysed as a present tense form. Elsewhere C. uses *got* as it it were a synonym of *have*:

(92) File 16
*CAR: I don't got a apple .

(93) File 16
*CAR: and I got two bears .

(94) File 14
*CAR: I got two lions .

Additional data presented in the following section further support the claim that there is no robust evidence to suggest that Agreement and Tense features have emerged in C.'s grammar by 3 years of age. The proportion of bare forms uninflected for Agreement and/or Tense clearly indicate that the child is still at the stage where no formal grammatical Agreement and Tense features have yet been discovered on the basis of the morpholexical input the child is exposed to.

7.6. Lexical verbs

English lexical verbs are assumed to carry weak Tense and Agreement features which block overt raising to T and AGR for checking purposes. Because -Interpretable Case and phi features must still be checked and erased in order to ensure convergence, the mechanism which allows checking is not Move but Attract. Tense and Agreement features on the lexical verb in V are attracted to T and AGR, they percolate to the FCs where they can be checked and, in the case of -Interpretable features, erased. In the adult grammar, matrix clauses require a finite verb which is specified for Tense and Agreement, while in child grammar many researchers have observed that children may optionally provide non-finite forms where the adult grammar requires finite forms. Whenever a non-target non-finite forms is present, some FC must be either not projected or underspecified (see sections 4.3.2 and 4.3.3 on various accounts of the Optional Infinitive phenomenon).

In this section C.'s performance on the obligatoriness of morphological finiteness markers is evaluated against competing accounts of the availability of FCs. It will be shown how a gradual shift from an initial almost complete absence of verbal forms, to a small number of bare forms and subsequently to a larger number of different verb forms testifies to the gradual and lexically-specific build up of grammatical marking encoded by lexical verbs. Even by the end of the recording sessions (3;1.25 ; MLUw 2.588) C.'s production of finite forms is still well below the acquisition threshold of 90% in obligatory contexts.

Because in English, finiteness in the present tense is only unambiguously marked on 3 p.s, these contexts were examined to investigate to what extent C. is consistent in marking finiteness in obligatory contexts. The identification of obligatory

contexts, however, is not always a straightforward task when it comes to child data. Even in the presence of a 3 p.s. subject, if the *-s* suffix is missing, one may not conclude with absolute certainty that the verb was intended by the child to be a present tense 3 p.s. form. An utterance like (95) below, may lend itself to a number of interpretations:

(95) Teddy fall over.

It could have a modal interpretation with a null modal auxiliary, *Teddy might fall over* (cf. Ingram & Thompson's (1996) null modal hypothesis), it could indeed be a valid context for 3 p.s. *-s* , *Teddy falls over*, it might be an imperative with Teddy being a vocative, *Teddy, fall over!*, or it could even be that the non-finite form is used *in lieu* of a past tense form, *Teddy fell over*. In order to try and establish with some degree of accuracy valid 3 p.s. contexts, only utterances where there was sufficient extralinguistic information to justify them were counted. Relevant information could either be the gestures accompanying C.'s utterance:

(96) File 2

*CAR: fall!
%exp: pushing the teddy bear off the armchair

or, the adult's recast of the child's previous utterance:

(97) File 2

*CAR: foot down.
*CAR: foot fall down.
*KAR: did his foot fall down?
*KAR: now what's gonna happen?
*CAR: fall!
*KAR: do you want the teddy to fall?

The results for present tense 3 p.s. contexts across the 19 English files are summed up in Table VI below:

File	+ *-s*	- *-s*	% in o.c.
1	-	-	-
2	-	-	-
3	-	-	-
4	-	-	-
5	-	-	-
6	-	-	-
7	-	3	0
8	-	-	-
9	-	-	-
10	-	-	-
11	-	-	-
12	2	2	50
13	-	-	-
14	-	-	-
15	-	-	-
16	-	2	0
17	5	10	33.3
18	-	15	0
19	3	1	75

Table VI. Present tense lexical verbs

The conservative estimate of 3 p.s. present tense contexts reveals that C.'s overall performance is rather poor, out of 44 obligatory contexts, only 8 verbs are correctly inflected (18.18%). In file 12 there are the first two tokens of inflected verbs, both with the verb *come*:

(98) File 12
*CAR: here comes the train .

(99) File 12
*CAR: no # this comes +...

The peak of 75% correct 3 p.s. *-s* in obligatory contexts in file 19 is a spurious figure given the extremely small number of tokens, moreover all of the three correct forms are instances of the verb *go*:

(100) File 19
*CAR: the giraffe goes here .

(101) File 19
*CAR: the cow goes like this .

(102) File 19
*CAR: the giraffe goes here .

In file 17 we find the most consistent use of 3 p.s. inflection, although here too the forms are all tokens of the same verb *go:*

(103) File 17
*CAR: this goes +...

(104) FIle 17
*CAR: it goes .

(105) File 17
*CAR: this goes there .

(106) File 17
*CAR: this goes there .

(107) File 17
*CAR: that goes there .

This is a clear example of lack of productive Agreement both in quantitative and qualitative terms. The percentage of correct forms in obligatory contexts is overall extremely low, less than 20%; furthermore, in those files where some use of 3 p.s. forms is indeed attested, only two verbs (*come* and *go*) ever appear in the inflected form.

As already seen in the case of the copula and the auxiliary be, and of the auxiliary *do*, all errors are errors of omission. There are no instances of overgeneralisation of 3 p.s. *-s* to non-3 p.s. contexts. The fact in itself is however hardly surprising since this particular suffix has still not been acquired by the child. It would be a very unexpected result if the child used *-s* in non-3 p.s. contexts before using it consistently in the required 3 p.s. contexts.

The results for past tense also show lack of control of Tense marking. Identifying obligatory contexts for past tense forms in spontaneous child data is as problematic as for 3 p.s. forms. In English there is the added complication that, at least

for regular verbs, the simple past and the past participle forms are homophonous, and in the presence of one such form it is not always easy to decide whether an *-ed* form is an auxiliaryless past participle or a target simple past form. Past-referring adverbials that could help in establishing fairly clear-cut contexts for past tense are unavailable in the data, therefore the only information available is context, as provided by adult utterances and recasts, and actions and/or events that can be clearly inferred to have taken place in the immediate or recent past (present perfect contexts), or in the more remote past (simple past contexts):[7]

(108) File 2

*KAR: what happened?
*CAR: fall down.
*CAR: fall down.
*KAR: what fell down?
*CAR: mouse.
*KAR: the mouse fell down.

(109) File 7

*CAR: fall dow(n).
*KAR: yeah it's fallen down.

(110) File 19

*CAR: and I cry.
*ERI: you cried?
*CAR: yes.

In both the example in file 2 and in the one in file 7, the event is the same: some object has just been knocked over and has fallen down. In both cases it is an event which has taken place in the immediate past. In file 2 the adult decides to use a simple past form twice (*what happened?* and *the mouse fell down*), while on a similar occasion the same adult uses a present perfect form (*yeah it's fallen down*). By contrast, in the third example only a simple past form is appropriate, the event being described (C. helping a friend at the nursery with a drawing) must have taken place a few days earlier, and as such is an instance of more remote past that is typically coded by simple past tense. Because of the ambiguity of obligatory contexts between simple past forms and present perfect forms in recent past events I have used the adult's utterances and recasts

[7]The perfective distinction between remote events and simple past, on the one hand, and more immediate past events, and present perfect on the other, is not a categorial one. The simple past is sometimes used in English even for very recent past events, therefore in counting obligatory contexts for either simple past or present perfect I have relied on the adult's utterances and recasts.

to determine Tense contexts. Table VII below sums up the results for simple past tense contexts, and Table VIII those for present perfect contexts.

File	+ *-ed*	- *-ed*	% in o.c.
1	-	-	-
2	1	5	16.66
3	-	-	-
4	-	-	-
5	-	-	-
6	-	-	-
7	-	5	0
8	-	-	-
9	-	-	-
10	-	-	-
11	-	-	-
12	1	1	50
13	-	-	-
14	-	-	-
15	-	-	-
16	1	-	100
17	-	-	-
18	6	-	100
19	4	5	44.44

Table VII. Simple past tense contexts

Out of 30 simple past tense contexts identified in the corpus, 13 are correctly supplied (43.33%). There are 8 different verb types (*fell, stopped, went, found, lost, said, wanted,* and *gave*), only 2 of which are regular verbs that take the *-ed* suffixation, the rest are suppletive irregular simple past tense forms, which may well not have been learnt as forms contrasting with a present tense form at all. For *found* in particular, there is reason to believe that C. does not in fact know that it is a past tense form, since on two previous occasions he uses it as the complement of a modal where *find* is required:

(111) File 15
*CAR: I can found them .

(112) File 15
*CAR: I can't found them !

For the other past tense forms, with the exception of *lost,* there are tokens of the bare form but no instances of 3 p.s., there is thus no strong evidence that the past

tense verbs in questions actually contrast with the corresponding present tense forms, but only with a bare uninflected form.

Table VIII presents the results for present perfect forms with or without the aspectual auxiliary *have*. Again, as in the case of simple past tense forms, obligatory contexts for present perfect were tabulated on the basis of either extralinguistic cues, or exploiting the discourse context set up by the adult's previous and subsequent utterances.

File	+ AUX	- AUX	% in o.c.
1	-	-	-
2	-	-	-
3	-	-	0
4	-	-	-
5	-	-	-
6	-	-	-
7	-	1	0
8	-	-	-
9	-	-	-
10	1	-	100
11	-	-	-
12	3	2	60
13	-	-	-
14	-	7	0
15	-	-	-
16	-	-	-
17	-	-	-
18	-	-	-
19	-	-	-

Table VIII. Present perfect contexts

Throughout the corpus there is only a very small number of present perfect contexts (14) out of which the auxiliary is correctly supplied only 4 times :

(113) File 10
*CAR: I've finished .

(114) File12
*CAR: oh no the train's stopped .

(115) File 12
*CAR: oh no it's fall down .

(116) File 12

*CAR: oh no it's fall down .

There are two examples in file 12 where the auxiliary is present but a bare form is supplied instead of the past participle, I have counted them as correct instances of present perfect as indicated by the presence of the auxiliary. The incorrect bare form could simply be considered as an instance of incorrect specific lexical knowledge whereby C. has not learnt that the past Participle of *fall* is actually the irregular *fallen.* With such a small data set it is difficult to draw any significant conclusions, it is however clear that some sort of avoidance strategy is in place and that C. privileges present tense contexts, and specifically present progressive contexts.

7.7. Discussion of C.'s use of verbs in English

When one looks at the overall distribution of verb forms in C.'s English corpus (403 in total), the picture that emerges is one where the largest number of verbs is made up of auxiliaryless present tense-*ing* forms (100), followed by an almost equally large proportion of present tense *-ing* forms correctly accompanied by an auxiliary (82), followed by modals (72), imperatives (66), bare forms (56), simple past tense forms (13), 3 p.s. present tense (10), and present perfect forms (4). If we discard imperatives, a large part of which is represented by the frozen form *look at that* (32/66, 48.48%), the only finite forms, as indicated by the use of an appropriate auxiliary, that reach any significant level are present tense progressive forms. Recall however that in section 7.3, where C.'s use of auxiliary *be* was analysed, it was emphasised how knowledge of inflected forms of *be* was essentially restricted to 3 p.s.. Moreover, out of 82 inflected forms of the auxiliary *be*, 12 (14.63%) are combined with the verb *doing* (11 of which are found in the semiformulaic question frame *what's x doing?*), and 40 (48.78%) with the verb *going*. These two verbs together account for 63.41% of all progressive forms combined with the inflected auxiliary *be.* It is a fact that C. uses these progressive forms with the auxiliary with correct subject-verb Agreement, however the high degree of lexical specificity in these progressive constructions is an indication that the child is still only gradually extending the *x is doing/going* pattern to a larger repertoire of lexical verbs.

It must be pointed out that the relative frequency of present tense progressive *-ing* forms may also in part be a sampling artifact. Given the nature of C.'s interaction in the free play sessions with the English-speaking adults it is almost inevitable that there

should be a large number of present progressive forms. When looking at picture books or playing with toy animals the adults tend to engage the child in conversation and elicit speech from him by asking questions about who the characters in the books are (naming games eliciting a large number of copula constructions of the type *that is a x*), or what the characters in the books are doing (e.g. *what's the bear doing here? who is the doggie talking to?*). The same is also partially true of more demanding activities such as playing with Lego or toy animals; there is a tendency for the adult-child dyad to give a running commentary on the actions being performed. When one looks at what kind of verb forms the adults interacting with C. use, it is therefore not surprising to find a disproportionate amount of progressive *-ing* forms vs. simple 3 p.s. present tense forms: 177 present tense *-ing* forms for E. vs. 80 simple present tense forms; 107 present tense -ing forms for K. vs. 31 simple present tense forms; 32 *-ing* forms for C. vs. 8 present tense, 28 *-ing* forms for S. and only 8 3 p.s. present tense. And finally J. who seems to be using more copula constructions than any other lexical verb: 7 tokens of 3 p.s. copula, 8 tokens of 3 p.p. copula, 5 *-ing* forms and 6 3 p.s. present tense forms.[8] Interestingly, in the adult data too, there is a correlation between certain verb types that occur in questions containing progressive *-ing* forms: for K. out of 26 *-ing* questions 20 are with the verb *doing*. For E. out of 31 *-ing* questions almost a third have either *going* (5) or *doing* (4) as predicate. C. only asks 7 *-ing* questions, 2 of which have a *doing* predicate. J. does not ask any *-ing* questions, and for S. the only *-ing* question she asks has a *going* predicate.

However, although some correlation exists between *-ing* questions and specific lexical items (namely *doing* and *going* in *what's x doing* and *where's x going*), nevertheless the number of verb types used by the adults in declarative *-ing* forms shows a lexical flexibility that is still absent in C.'s *-ing* forms.[9]

---

[8] See Table IV in section 3.7.2 for a list of participants in the English sessions.

[9] Table A below presents the number of progressive verb types and tokens used by the adult participants in interaction with C:

Adult Participant	*-ing* verb types	*-ing* verb tokens
K.	31	107
E.	15	177
C.	17	32
S.	11	28
J.	4	5

TableA. Distribution of -ing verb types and tokes across the adult participants.

As for overlap between different inflected forms of the same verb type, the only two verbs that appear with 3 p.s. inflection (*come* and *go*) also appear with *-ing* ending, and for *go* one instance of simple past *went* is also attested in file 14:

(117) File 14
*CAR: she went clup@o .

Of the 8 simple past tense forms (*fell, stopped, went, found, lost, said, wanted, gave*), apart from *go*, none have a corresponding 3 p.s. form, and only 2 (*fell* and *went)* also appear in the progressive *-ing*. *Go* is the only verb that appears with all of the 3 inflections: *-ing*, *-s,* and *-ed,* although the past form is a suppletive irregular form, and because it only appears once, in a somewhat atypical meaning, it is not at all clear to what extent it is actually paradigmatically related to *go* and *going*.

There is no direct positive evidence for a productive mechanism of verb-general marking in C.'s English data. The only productive morphological rule seems to be *-ing* suffixation on eventive predicates. These observations, combined with the fact that there is very little evidence for productive subject-verb Agreement in the present tense paradigm of lexical verbs, show that by age 3;1 there is yet no strong indication that Agreement and Tense are part of the child's grammar.

On the other hand, the aspectual notion of imperfectivity as encoded by *-ing* morphology is available from file 8 (2;2.4) onwards when C. starts using -ing forms. Form file 12 some Tense and Agreement features start to emerge whereby *-ing* forms are accompanied by an appropriately inflected auxiliary an average of 70% of the time. It is as if lexically Aspect-marked *-ing* forms triggered obligatory temporal fixation through the specification of an auxiliary inflected for Tense and Agreement.

Judging by the predicates that appear in the -ing forms (*hiding, doing, crossing, going, coming, bringing, falling, crying, swimming, getting, putting, making, playing, eating, raining*), C. is also sensitive to the eventive/stative distinction whereby only eventive predicates can take imperfective progressive aspectual *-ing* morphology. As remarked by Brown (1973), progressive forms in early English are consistently used with non-stative verbs from the very beginnig. Overgeneralisation, yielding ungrammatical sentences such as *I am knowing you*, where a stative verb like *know* is suffixed with *-ing*, is only attested once:

(118) File 17

*CAR: it's being like a person.

In the spirit of Tsimpli's (1996) analysis of Aspect-marked forms in early child English, it is argued here that what the child is initially sensitive to is the perfective/imperfective contrast and progressive/habitual distinction as expressed by the language's morphology. The *-ed* ending specifies both perfective Aspect and past tense, while present tense forms and *-ing* forms are specified as imperfective. Present tense and -ing forms are distinct with respect to the progressive feature, with present tense forms being specified as [-progressive] and -ing forms as [+progressive]. Because of the lack of salient Agreement or Tense morphology, virtually non-existent in the present tense, and conflated with Aspect morphology in the past tense, the only salient contrast that English makes explicit is an aspectual one. The child will therefore initially focus on those salient aspects that are thus incorporated in his/her grammar. The relevance of Tense and Agreement features will gradually become available when the child notices that these features are marked with some consistency only on *be,* where the child can appreciate Person contrast in the singular half of the paradigm (3 different forms for the 3 different Persons), and Number contrast between the singular Persons and the three plural Persons (indistinguishable from one another, but contrapposed to 1p.s. and 3 p.s.).

Although C.'s performance on verbs is not consistent with the existence of Agreement and Tense features that require movement to AGR and T categories to be checked, nevertheless his performance is surprisingly good on obligatory subjects and specifically on subjects' Nominative Case marking. The next section investigates this phenomenon in greater detail to assess its significance and its consequences for an accurate description of the child's grammar.

7.8. Overt and null subjects

The standard assumption is that Nominative Case is checked in spec-AGRS (cf. Ingham, 1998, Watanabe, 1996). It follows that if an AGRS projection is not present in the child's grammar, subjects should in principle not be in a position to check their Nominative Case and should therefore surface with default Case, which in English happens to be accusative Case (cf. Schütze & Wexler, 1996). Moreover, the requirement of obligatory overt subjects in English is a language-specific option that is still poorly understood, but that seems to be related to the morphological richness of

the AGRS node. In null-subject languages such as Italian and Spanish, a rich AGRS node allows identification of the content of *pro* subjects, INFL-licensed null subejcts (Hyams, 1986; Rizzi, 1994). Null subjects are also available in discourse-oriented languages like Japanese where verbs do not carry Agreeement features and an AGRS node is thought not to exist at all, and where null subjects are identified via coindexation with an appropriate discourse antecedent (Huang, 1984). A number of studies investigating the production of subjects have found a correlation between finite forms and overt subjects on the one hand, and root infinitives and null subjects on the other (Sano & Hyams, 1994; Hyams, 1996 for English; Krämer, 1993 for French and Flemish; Hamann, 1996 for German; Haegeman, 1994 for Dutch, Hamann & Plunkett, 1996 for Danish, Krämer, 1993 for Flemish). Finite verbs tend to occur with overt subjects, while RIs tend to occur with null subjects or indeterminate DPs which are not specified for Number (Hoekstra, Hyams & Becker, 1996).

This section will investigate the production of obligatory subjects in C.'s English data. Table IX below sums up the results for the total number of subjects with finite and non-finite verbs.

File	+subj	-subj	% subj in o.c.
1	-	-	-
2	2	3	40
3	2	-	100
4	-	-	-
5	-	-	-
6	-	-	-
7	3	4	42.85
8	-	2	0
9	10	-	100
10	24	-	100
11	33	2	94.28
12	55	8	87.30
13	51	5	91.07
14	102	9	91.89
15	52	5	91.22
16	46	4	92
17	44	7	86.27
18	79	3	96.34
19	63	3	95.45

Table IX. Total number of overt and null subjects

What is immediately clear from the results in Table IX is that as soon as verbs start appearing with some consistency, from file 9-10 onwards, the production of overt subjects in obligatory contexts is extremely high. This positive correlation between verb

use and subject use is also reported by Valian (1991) for a group of American English-speaking children (age range 2;0-2;5; MLU 1.77-3.39), and in Valian & Eisenberg (1996) for a group of Brazilian Portuguese-speaking children (age-range 2;5-2;6; MLU 1.71-3.83).[10] The occurrence of verbs is a very good predictor of the occurrence of subjects in C.'s data. In fact he does not seem to go through the well-attested Optional Subject Phase that characterises the early linguistic stages of children acquiring non-null subject languages. Nevertheless, there is still a proportion, albeit small (56/637, 8.79%), of ungrammatical null subjects.

In Table X the distribution of overt and null subjects is presented for 7 different categories of verbs: bare forms, 3 p.s. simple present tense forms, progressive present tense forms, past tense forms, modals, copula, and others. The "others" category include verbs which cannot unambiguously be classed as either finite or non-finite, that is verbs with a non 3 p.s. subject:

(119) File 19
*CAR: I go and close the door .

Although the utterance in (119) above could be considered as a grammatical present tense utterance with a 1 p.s. Nominative subject, nevertheless, because in English non 3 p.s. lexical verb forms are indistinguishable from a bare non-finite form, it is not possible to conclude with absolute certainty that these are truly examples of finite verb forms. They have however been included in the sample because they account for a substantial part of overt subjects.

[10]The Italian children also included in Valian's (1991) study failed to show an increase in the number of subjects over the one-year period that they were recorded Valian & Eisenberg (1996) suggest that this result could simply be a sampling artifact. Because the children in the Italian group were younger than 2;6 when the study ended (the average age of the Portuguese-speaking children) it cannot be excluded that an increase in subject use could have been observed in the Italian group too, had the study lasted longer.

File	bare	3 p.s.	Past	Prog.	modals	copula	neg	other	% o.s.
1	-	-	-	-	-	-	-	-	-
2	2/3	-	-	-	-	-	-	-	40
3	-	-	-	-	-	2	-	-	100
4	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-
7	3/4	-	-	-	-	-	-	-	42.85
8	-/2	-	-	-	-	-	-	-	0
9	-	-	-	-	-	8/-	-	2/-	100
10	-	-	-	-	-	18/-	1/-	6/-	100
11	-/1	-	-	-/1	-	4/-	3/-	9/-	94.28
12	2/2	2/-	1/-	17/2	-	27/3	3/-	5/-	87.3
13	1/-	-	-	-	22/1	3/2	1/1	25/2	91.07
14	3/-	-	7	17/3	1/-	16/5	11/-	41/1	91.89
15	3/-	-	-	18/2	6/-	18/2	-	7/4	91.22
16	6/-	-/1	2/-	4/1	1/-	22/2	2/1	12/1	92
17	11/1	6/-	-	6/1	1/-	15/3	2/-	4/-	86.27
18	18/1	-	7/-	4/1	21/-	22/-	1/2	7/1	96.34
19	5/-	3/-	3/-	4/-	14/-	16/-	1/-	17/-	95.45
Total	54/14	11/1	20/0	52/11	66/1	169/17	25/4	109/9	
% o.s.	79.41	91.66	100	80.95	98.50	90.86	86.20	92.37	

Table X. Distribution of overt and null subjects by verb category.[11]

The lowest percentage of overt subjects is with bare forms, this in itself is not a surprising result. This correlation between RIs and null subject has been previously documented in a number of studies in different non-null-subject languages. What is instead very surprising, is that the percentage of null subjects with RIs is extremely low in C.'s data compared with previous findings in the literature. In a review of the literature on RIs and null subjects, Hoekstra, Hyams & Becker (1996) report that, crosslinguistically, RIs typically occur with null subjects roughly 83% of the time. C.'s data goes exactly in the opposite direction, almost 80% of RIs have an overt subject. How is one to account for this puzzling fact? Hoekstra et al. (1997) argue that not all overt subjects are alike, some lexical subjects are finite and some are non-finite. Finite lexical subjects are determinate DPs, i.e. DPs with a determiner; non-finite subjects can be null (PRO subjects), or lexical indeterminate DPs (bare singular nouns). At the basis of the non-specificity of these indeterminate DPs is the underspecification of Number feature (cf. section 2.3.2), therefore only default singular nouns are expected. Because RIs are also the result of the underspecification of Number, they are predicted to be more likely to occur with non-finite lexical subjects which are equally underspecified for Number. Indeed, when one looks at the type of subjects that appear with RIs in C.'s data, there is some evidence that specificity may well play a role. Out of

[11]The figures to the left of the slash represent the number of overt subjects, while those to the right are for the null subjects. The last row (% o.s.) stands for % of overt subjects.

a total of 54 overt subjects, 24 are demonstrative pronouns *this* and *that,* 2 are bare nouns, 7 are proper names, 14 are personal pronouns (*I, he/she, we*), and 7 are DPs containing a determiner, one of which includes a plural noun (*the squirrels*).

The demonstrative pronouns *this* and *that* are clearly all-purpose placeholders that are not specified for either number of definiteness; together with the 2 bare nouns they make up for 48.14% of all overt subjects found with RIs. The remaining 51.86% of RIs' overt subjects are personal pronouns (25.92%), proper names (12.96%), and determinate DPs (12.96%).

Hoekstra et al. (1996: 299) discard tokens of proper names and pronouns from their count because "they can occur with both finite and non-finite verbs". Although they do not further qualify the rationale behind this statement, following Longobardi (1994), one could argue that proper names in English are base-generated in the N position inside the DP, and do not raise overtly in the syntax (see chapter 5). Pronouns, by contrast, are supposed to be generated under D, and are assumed to be inherently specific, it is therefore more difficult to see why Hoekstra et al. would want to allow pronouns to appear as RI subjects, when there would be a feature mismatch between the underspecification of Number in the verbal domain and the specification of Number in the nominal domain. Pronouns are therefore to be considered as determinate DPs, and together with the 7 DPs containing an overt determiner they account for 38.88% of all overt subject with RIs.

Although the majority of RI subjects is still made up by *this/that* subjects and bare singular nouns (48.14%), nevertheless a non-negligiable proportion of RI subjects are indeed specified for Number, an unexpected finding for the Underspecification of Number Hypothesis. Even in the Hoekstra et al.'s study, only data from one of the two English children examined support the original prediction of a correlation between RIs and indeterminate DPs. Out of 41 subject DPs with RIs, only 2 have an overt determiner in Adam's data, while the remaining 39 have a null determiner (data available in the CHILDES database, MacWhinney & Snow, 1985). DP subjects of finite verbs contain an overt determiner 53 times out of 57. In Adam's case the asymmetry between finite and non-finite forms with respect to the specificity of subject DPs is very clear-cut and goes exactly in the direction predicted by Hoekstra and colleagues. In the case of Nina, the other child whose data is reported in the study (data available in the CHILDES database, MacWhinney & Snow, 1985), the correlation is not so obvious, in fact she has a greater proportion of overt determiners both in RI subjects (12/21) and in

finite verbs (34/37). The proposed solution is to argue that these determinate DPs are not really subjects, but rather dislocated constituents. Although dislocation of a DP through topicalisation is an option that is allowed in the adult grammar of English, it is a marked phenomenon which, incidentally, also involves a resumptive pronoun:

(120)a. Simenon$_i$, I don't like him$_i$.

b. Miss Marple$_i$, she$_i$ was always my favourite.

Hoekstra et al. do not provide examples of the alleged dislocated constituents in Nina's corpus, therefore it is not possible to say anything more about these DPs. One can however speculate that, if they were really examples of dislocated constituents, they should be coindexed with an appropriate resumptive pronoun. In C.'s data this is certainly not the case. It could still be maintained that there is in fact a phonetically null resumptive pronoun *in lieu* of the overtly realised adult counterpart, a somewhat difficult position to maintain, if there are never any examples of left-dislocated DPs with a resumptive pronoun in the data. A second problem is that out of C.'s 21 determinate DPs, 14 are personal pronouns that could not be dislocated anyway, a fact which further weakens the dislocation hypothesis.

An alternative explanation of C.'s pervasive use of subjects, both with finite and non-finite verbs, would be to argue that the child attends to the obligatoriness of subjects in English as an independent grammatical property acquired in its own right, independent of the acquisition of the notion of subject-verb agreement in terms of feature checking. Apart from present tense 3 p.s. contexts and copula and auxiliary *be* and *do* contexts, where matching of nominal and verbal Person and Number Agreement phi features is required, in all the other contexts (past tense contexts, modal contexts, non 3 p.s. present tense contexts, RI contexts) any subject with any Person and Number combination can be selected, the verb is unmarked with respect to these phi-features, hence no mismatching can occur. The number of cases where phi feature matching between subject and verb is required is 261 out of 563 (46.35 %), in all cases where an overt subject is produced it does agree with its verb, a *prima facie* indication that C. must know about Agreement pheatures and feature checking. It has however been amply demonstrated in the previous sections of this chapter that there are principled reasons to question this conclusion. If subjects are not actually selected because of the Agreement features they spell out, then one could speculate that the reason they are produced at all, even with a very large proportion of bare forms where a

PRO subject would be perfectly well justified, is that C. appreciates that subjects are required in English.

Hyams (1986) has made the suggestion that there may be a pragmatic component to the emergence of obligatory subjects in English. One of the triggers for the expression of overt subjects in English may be the recognition by the child that subjects occur in the input even when they are communicatively redundant. Ingham (1992) makes a similar proposal for the disappearance of the Optional Subject Phenomenon in the data of Sophie, a monolingual English child observed between the ages of 2;5 and 2;8. Ingham reports that by the time Sophie produced subjects in 90% of obligatory contexts, she still had not acquired any of the Agreement and Tense features that are typically associated with the emergence of AGR and T. Although the mastery of inflection is normally assumed to be one of the pre-requisite for obligatory subjects in non-null subject languages, it may well be that there is a non-negligiable pragmatic component to the acquisition of a syntactic property that in fact has not even been explained all that satisfactorily by linguistic theory.

7.9. Conclusions

A fine-grained analysis of C.'s use of verbs in English has revealed that the child only ever commits omission errors, there are no commission errors in his use of morphologically inflected forms. He is a very conservative learner whose main strategy is to rely on a comparatively small number of positional patterns, and a limited repertoire of verbs that he tends to use with a small variety of subjects and complements. By the end of the period of observation (3;1.25) there are no conclusive indications that Agreement and Tense features are part of the child's grammary yet. In one respect, however, his performance is exceptionally good: production of obligatory subjects. This achievement however seems to have a pragmatic basis, more than a syntactic one, given the overall poor performance in other areas of the grammar that are supposedly tied to the emergence of subjects.

Chapter 8

Strategies in the acquisition of Italian morphology

8.1. Introduction

In chapter 7 C.'s performance on English verbs was analysed and it was concluded that although there is a gradual progress in the use of the copula and of the auxiliary *be*, and some evidence of correct use of modals, there is no clear indication that Agreement or Tense morphosyntactic markers are being productively used by the end of the period of observation (3;0.17). The degree of lexical specificity with which verbal inflections are correctly used with a very small set of verbs in a limited number of semiproductive positional patterns suggests, at the very least, that C. cannot be safely credited with across-the-board competence in the use of these inflections. Rather his knowledge seems to be still very much tied to the lexical items with which such Agreement and Tense inflections are used. Although the possibility exists that the small scale production of these morphosyntactic devices is due to a sampling artifact, there is reason to believe that this cannot be the only explanation. A comparison between the number of verb types used by C. and his adult interlocutors shows that overall the adults use a larger repertoire of verbs than the child does. It is not simply the case that the situation necessarily requires the use of such a small number of verbs as C. uses.

In English, C. displays the behaviour of a very conservative learner, his errors are overwhelmingly omission errors and not commission errors. His strategy in English is to rely on a number of well-rehearsed patterns which can be safely used again and again with minimal morphological, syntactic and lexical variations.

In this chapter, C.'s performance in Italian will be analysed in detail. It will be shown that lexical specificity also plays a role in C.'s Italian, although the number and variety of inflected forms is larger in absolute terms than in English. Section 8.2 will look at the use of the copula *essere*, section 8.3 will present data on the use of lexical verbs and their productive and contrastive use over time, while section 8.5 will draw some conclusions.

8.2. The Italian copula

8.2.1. Production in obligatory contexts

The paradigms of the present tense and imperfect tense Italian copula *essere* are given in Table I and II below:

1 p.s	sono
2 p.s.	sei
3 p.s.	è
1 p.p.	siamo
2 p.p.	siete
3 p.p.	sono

Table I. Present tense copula *essere*

1 p.s	ero
2 p.s.	eri
3 p.s.	era
1 p.p.	eravamo
2 p.p.	eravate
3 p.p.	erano

Table II. Imperfect tense copula *essere*

Apart from the present tense paradigm where there are two homophonous forms (1 p.s. and 3 p.p.), all the other forms are distinct for the different singular and plural person/number combinations. For the present tense in particular, all the forms are highly suppletive and in no obvious way related to the infinitival form *essere*.

In order to assess C.'s use of the Italian copula, the corpus was automatically searched for tokens of present tense, imperfect and present perfect tense copula *essere*; in addition, a search for obligatory copula contexts was also performed automatically and disambiguated manually. Potential obligatory contexts include all those predicative constructions containing a null or overt subject (be it a pronoun, a proper noun, or a common noun), and a nominal or adjectival predicate of the type *DP + essere + DP*. The results for present tense and imperfect tense contexts are shown in Table III and VII respectively, no contexts for present perfect were found. In both tables, the figure before the forward slash "/" indicates the number of tokens for a given person, and the figure after the forward slash expresses the total sum of obligatory contexts (i.e. the number of the tokens produced for that given person and

the number of contexts where the copula for that particular context is missing). The figure before the slash includes both correct and incorrect tokens for the given person.[1]

File	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.	% in o.c.	MPHs
1	-	-	-	-	-	-	-	-
2	-	-	3/3	-	-	-	100	9
3	-	-	5/7	-	-	-	71.42	30
4	-	-	6/11	-	-	-	54.54	4
5	-	-	0/1	-	-	-	0	-
6	-	-	16/24	-	-	-	66.66	19
7	-	-	66/96	-	-	0/2	67.34	3
8	-	-	12/13	-	-	-	92.3	-
9	-	-	14/16	-	-	-	87.5	21
10	-	-	35/37	-	-	-	94.59	-
11	-	-	17/17	-	-	1/3	90	-
12	-	-	16/18	-	-	-	88.88	-
13	-	-	28/28	-	-	1/1	100	-
14	-	-	64/64	-	-	5/6	98.57	-
15	2/2	-	41/42	-	-	1/1	97.77	-
16	-	-	14/14	-	-	5/5	100	-
17	-	-	39/39	-	-	10/10	100	-
18	-	-	14/14	-	-	5/5	100	-
19	-	-	12/12	1/1	-	10/10	100	-
20	-	-	23/23	-	-	8/8	100	-

Table III. Present tense copula *essere* in obligatory contexts

The eighth column gives a total percentage of copula production in obligatory contexts, while the ninth column indicates, where this is relevant, the number of Monosyllabic Place Holders (MPHs) that appear in each file. Monosyllabic Place Holders, as defined by Bottari, Cipriani & Chilosi (1993/94), are [-tense] vocalic segments that perform the function of protomorphemes. For convenience Bottari et al.'s term (MPHs) is adopted here, however it will be made clear in the presentation of the results below, that the interpretation given in this work differs somewhat from Bottari et al.'s original proposal.

As was already observed for the English copula (see section 7.2), the only two contexts that are represented with any consistency in the data are 3 p.s., and, to a lesser extent, 3 p.p.. Once again, the large number of 3 p. contexts is related to the kind of activity C. is engaged in with his adult interlocutor. During the recording sessions some kind of prop is normally used to elicit speech from C. (see Table III in section 3.7.2 for a complete list of activities by recording session). Looking at picture books, playing with toy animals, and drawing are three of the activities that feature most

[1] This first measure of the copula's productivity is simply intended to evaluate C.'s sensitivity to the presence of this grammatical marker. Section 8.2.2. presents a more fine-grained analysis of commission errors with copula forms and their implications.

prominently throughout all of the data collection period. It is therefore inevitable that much of the verbal exchanges between C. and the adult will have the characters in the books, the toys, or the drawings the child does, as the joint focus of attention and topic of conversation, hence the large number of 3 p. contexts found in the sample.

The copula is found in a number of constructions: *null subject + copula + x*, where *x* can be an adjective or a noun ((1)a and (1)b); *demonstrative pronoun + copula + x*, where *x* is either an adjective or a noun ((1)c and (1)d); *existential ci + copula + noun* ((1)e); *null subject + copula + locative pronoun* ((1)f):

(1)a File 4

*CAR: è un lupo.
%eng: it's a wolf.

b File 14

*CAR: sono nere.
%eng: they're black.

c File 7

*CAR: questa è (l)a mamma.
%eng: this is the mummy

d File 7

*CAR: anche questo è blu.
%eng: this too is blue.

e File 9

*CAR: c'è la volpe.
%eng: there is the fox.

f File 6

*CAR: è lì.
%eng: it is there.

8.2.2. The significance of low error rates

From file 8 (2;1.23, MLUw 1.904) onwards, the 3 p.s. copula is consistently attested in around 90% and over of obligatory contexts, reaching a stable 100% from file 13 onwards (2;5.6). As already mentioned above, the raw numbers include both correct and incorrect tokens of 3 p.s., i.e. both cases in which the 3 p.s. copula

correctly agrees with a 3 p.s. subject, and cases in which the 3 p.s. is inappropriately used with an overt plural subject.

As for the 3 p.p. copula, although it does not make its first appearance until file 11 (2;3.7) and the absolute number of tokens totals up a mere 46 vs. 425 for 3 p.s., nevertheless it reaches the same overall level of suppliance as 3 p.s.: a total of 425 tokens of 3 p.s. out of 479 obligatory contexts (88.72%), vs. 46 tokens of 3 p.p. out of 51 obligatory contexts (90.19%). A number of other studies also report that, initially, plural verb forms are produced more rarely than singular forms in Italian (Hyams, 1986, 1992; Pizzuto & Caselli, 1992; Guasti, 1993/94). Some researchers (Guasti, 1993/94, Hyams, 1992, Hoekstra, Hyams & Becker, 1996) have made the suggestion that the smaller number of plural verb forms is principally due to a manifestation of the general phenomenon of late development of plurality. One important consequence which follows from this argument is that, if the smaller number of plural forms is simply the reflex of some pragmatic "avoid plural phenomenon" (cf. Hoekstra et al., 1996), and the child has in fact productive knowledge of the notion of subject-verb agreement, one would expect to find comparable error rates in the use of singular and plural forms (cf. Rubino & Pine, 1998). If, on the contrary, different error rates are observed for singular and plural forms, it is more difficult to claim that subject-verb agreement is a notion that applies across the board, rather than the outcome of verb-specific piecemeal learning (cf. Gathercole, Sebastian & Soto, 1999).

The data from Table III is further broken down for correct and incorrect agreement in Tables IV and V, where the extent of agreement errors is clearly indicated. Only one specification error was found. An agreement error is one where the subject and the verb do not agree in person and/or number, as for example in (2):

(2) File 7, line

*CAR: c'è i gufi.
%eng: there is the owls.

A specification error occurs when agreement between subject and verb is correct, but the specification of the referent is incorrect and for example the agreement is in the singular number when the referent is in fact plural:

(3) File 7, line

*CAR: è rossa.
%eng: it is red.

In the example in (3) C. correctly uses a 3 p.s. singular copula and a singular feminine adjective, but the referent (*ruote, "wheels"*) is plural and a plural copula and a plural adjectives would be required instead.

Table IV presents the total number of copula strings in the Italian sample, together with the relative proportion of singular and plural forms:

COPULA FORMS	N	%
singular copula forms	424	90.02
plural copula forms	47	9.98
total production of copula forms	471	100

Table IV. Total production of singular and plural copula forms

As already mentioned above, plural forms are a minority of the overall number of copula tokens, they account for less than 10%. Table V presents the results for correct and incorrect utterances involving subject-verb agreement with present tense copula *essere* in C.'s Italian sample:

AGREEMENT	N	%
correct subj-copula agreement	463	98.3
incorrect subj-copula agreement	8	1.7
total production of subj-copula	471	100

Table V. Overall production of correct and incorrect copula strings.

As can be seen from Table V above, the overall error rate is extremely low, a mere 1.7% of the total number of utterances containing a form of the copula shows incorrect agreement. It is now of interest to see more in detail the distribution of agreement errors as reported in Table VI:

SUBJECTS	correct agreement		incorrect agreement	
	N	%	N	%
singular subject	416	100	0	0
plural subject	47	85.45	8	14.55

Table VI. Distribution of correct and incorrect agreement with singular and plural subjects

When collapsed across singular and plural forms the overall error rate for the present tense copula is extremely low (1.7%); however, when one looks at the

distribution of the incorrect and correct forms over singular and plural forms, it becomes apparent that incorrect agreement is only ever found with plural subjects taking a singular form of the copula, and in particular with 3 p.p. subjects taking a 3 p.s. copula:

(4) File 10

*CAR: c'è i pinguini.
%eng: there's the penguins.

A more fine-grained analysis of the error rate in copula constructions reveals an asymmetry between the mastery of the singular and the plural paradigm. The fact that tokens of the plural copula are rarer, does not necessarily imply that in principle the error rate should be any different from what it is for singular instances of the copula. In fact, if subject-verb agreement is a mechanism which is productive across the board, one would expect comparable error rates between singular and plural forms, which is not the case in C.'s data. *Prima facie*, it would be reasonable to conclude from the results in Table I and Table VI that realisation of the obligatoriness of the copula is a gradual process which begins to be attested with some consistency for 3 p.s. forms around file 8 onwards (2;1.23), while 3 p.p. forms begin to emerge in file 13 (2;5.6) but a number of incorrect forms persist as late as file 20 (3;0.17, MLUw 3.306):

(5) File 20

*CAR: cos'è spari?
%eng: what is shots?[2]

It is important to observe that out of 8 incorrect form with a plural subject and a singular copula, 5 are of the form *c'è + plural DP*:

(6)a File 7

*CAR: c'è i gufi.
%eng: there is the owls.

b File 10

*CAR: c'è le pecore.
%eng: there is the sheep.

[2]In the example in (5) the plural noun spari, "shots", could be a quotative, and therefore it would not be treated by the child as intrinsically plural. If this is the case the lack of agreement would be justified.

c File 10

*CAR: c'è i pinguini.
%eng: there is the penguins.

d File 10

*CAR: c'è le scarpe.
%eng: there is the shoes.

e File 15

*CAR: non c'è più animali.
%eng: there is no more animals.

The recurrent pattern in 62.5% of the errors with a plural subject and a singular copula in the existential constructions in (6) can be explained if one thinks of C. starting out by rote-learning the *c'è + DP* construction where *c'è* is analysed as one invariable element in construction-initial position typically followed by a singular DP, but on occasion also by a plural DP. It looks as if, at least in part, C. is relying on some semiformulaic utterances where *ci* is part of the frozen chunk *c'è* which can be followed by any DP, regardless of its number specification. In this sense the string *c'è* behaves as an existential placeholder which is not specified for Person or Number features, as suggested by Groat (1995) for English existential *there*, and as such can be followed by any DP with any combination of person and number features.

8.2.3. MPHs in C.'s copular constructions

The idea that the monosyllabic string *c'è* could function as a placeholder is particularly relevant for the analysis of a set of data found in the earliest of the Italian files in C.'s corpus. Column nine in Table III reports a number of what have been defined as Monosyllabic Place Holders (MPHs), monosyllabic elements that replace either whole word clusters, or function as protomorphemes. Bottari et al. (1993/94) identify two developmental stages in the distribution of MPHs. An initial stage is characterised by the occurrence of [-tense], [-round] vowels, sometimes with a nasal feature added. The range of forms is subject to variation, not only between different children, but also in the output of an individual child. A number of examples taken from Bottari et al. (1993/94: 331) are provided in (7):

(7)a. [ˆ] [ˆ] miao
[ˆ] [mjao] child form for "cat"

b. [e] [e] apre
[e] it-opens

c. [O] [O] divano
[O] sofa

d. [A] [A] pappa
[A] mush

e. [a] [a] zia
[a] aunt

f. [U] [U] bimbo
[U] child

The authors observe that MPHs are not associated with a specific phonetic feature, but rather with a phonetic space, the low-central area which typically contains "unmarked" segments. In this sense, MPHs lack a clear phonological status and cannot be considered as phonetic approximations to specific morphemes, but should rather be viewed as underspecified fillers of structural positions. According to Bottari et al., MPHs begin to emerge in a period when MLUw ranges from 1.2 to 1.6, and when approximately 80% of utterances are still one-word utterances. At this stage MPHs cooccur in the same string with a conventional word which can be a referential item, an adverb or a pivotal form (examples taken form Bottari et al., 1993/94: 332):

(8)a. [e] bimbo
[e] child

b. [e] basta
[e] enough

c. [a] lì
[a] there

d. [a] via
[a] allgone

At this early stage, the interpretation of MPHs can be ambiguous, and in certain circumstances it can be inferred that they do not in fact stand for one single free morpheme, as for example a determiner before a noun, but it may well be the case that they actually replace a whole string of words. In (11)a for instance [e] could stand for *il "the", c'è un/il, "there is a/the", guarda il "look at the",* etc. At a time when the child's utterances are still very short and simple, inferring the intended meaning is clearly a difficult task. Familiarity with the child, and with the specific situation, in

addition to access to the extralinguistic context is often necessary to establish with some degree of accuracy what these MPHs stand for in the child's utterances.

In a second stage, where MLUw is higher than 1.6, and two-word and multiword utterances start being produced, the overall proportion of MPHs increases and their distribution changes significantly. Their interpretation is virtually always unequivocal, and, unlike in the earlier stage where they could be seen as replacing open-class words such as nouns and verbs, now they unambiguously stand in for closed-class elements such as articles, prepositions, clitics, copula, negative and interrogative operators:

(9)a. torna [a] (= il) babbo.
he-comes back [a] (= the) daddy

b. io [e] (= nella) ciabbia.
I [e] (= in the) sand.

c. [e] (= me) prende
[e] (= me) he-takes

d. [ˆ] (= è) fori.
[ˆ] (= it's) outside

e. [e] (= non) c'entra
[e] (= not) there-enters

f. [ˆ] (= che) fai lì?
[ˆ] (= what] you-do there?

An analysis of the MPHs that appear in C.'s Italian data shows that they belong with the MPHs that Bottari et al. identify for the earlier stage. A first observation that leads to this conclusion is that MPHs occur between file 2 and file 9 where the MLUw ranges from 1.165 to 1.883, a MLUw stage comparable to that proposed by Bottari et al. for the earlier stage. Secondly, the interpretation of C.'s MPHs clearly indicates that they always stand for word clusters and not simply for closed-class free morphemes. More specifically, all of C.'s MPHs are found in matrix questions and replace a sequence of the type *wh-word + (existential ci) + 3 p.s. copula.* MPHs are found in combination with a noun optionally preceded by a determiner, an adverb, or a demonstrative pronoun:

(10)a. File 3

*CAR: [^] (= dov'è) cane?
%eng: [^] (= where is) dog?

b. File 6

*CAR: [^] (= cosa c'è) dentro?
%eng: [^] (= what is there) inside?

c. File 7

*CAR: [^] (= chi/cos'è) questo?
%eng: [^] (= who/what is) this?

The repertoire of C.'s MPHs includes [ ], and also the syllable [de], although in a smaller proportion and mainly found in file 9 (17 tokens of *[de] afante?,"where elephant?"*):

(11) File 9

*CAR: [de] (= dov'è) afante (= elefante)?
%eng: [de] (= where is) elephant?

The interpretation of the MPHs in (10) and (11) is particularly unambiguous in that the questions they are embedded in are clearly modelled on the adult input, and are part of a familiar interactional routine. The question *dov'è x?,"where is x?"*, is normally asked by the adult when looking at one of C.'s favourite picture books containing flaps that can be lifted to reveal an animal hiding. After the adult's prompt, usually *dov'è x? ,"where is x?"*, but also *chi/cosa c'è dentro? ,"who/what is there inside?"*, or *cos'è questo? ,"what is this?"*, C. is expected to reply either by lifting the flap and/or saying *è qui ,"it's here"*, or by providing a name, e.g. *l'orso/c'è l'orso/è un orso ,"the bear/there's the bear/it's a bear"*.

(12) File 7

*LUD: chi c'è qui dentro?
%eng: who is there inside here?
*CAR: il leone.
%eng: the lion.

In the examples in (10) and (11) C. has taken on the adult's role and he is now delivering the adult's lines of a well-rehearsed script. This role reversal typically takes place when the child has become familiar with the routine and is confident enough not only about his own part, but also the adult's (cf. Peters & Boggs, 1986). Given the

unambiguous source of C.'s MPHs, it is therefore clear that they replace a sequence of words, *wh-word + (existential ci) + copula*, and not simply some functional morpheme as suggested by Bottari et al. (1993/94) for some of the later MPHs found in their corpus. These do not replace content words or open class items, but occur in contexts where a closed-class word would be expected. They hold the place of various free functional morphemes, some of which have made their first appearance but are not yet used productively.

The fact that in C.'s corpus there are only MPHs that replace whole phrases, as opposed to MPHs that stand in for free functional morphemes, suggests that the child is treating these sequences of *wh-word + (existential ci) + copula* embedded in the adult's wh-questions more like one single unit than as individual lexical items. This is also consistent with analogous copula constructions in C.'s English (see section 7.2). In English the copula appears in interrogative contexts in one of two frozen forms: *what's that?* (4 tokens) and *where is it?* (13 tokens). As for declarative contexts there are only two constructions where the copula is found: *that's (a) x* and *it's (a) x*. In Italian, as has been shown above, the copula appears in a wider number of contexts, although the subject position in copula constructions is restricted to null subjects, demonstrative pronouns, and existential *ci*. There are no Italian copula constructions in C.'s data in which the subject position is filled by a DP containing a determiner + noun, a proper noun, or a pronoun. Even though in Italian the copula does appear in a wider number of contexts, this simply shows that C. has a larger repertoire of copula constructions in Italian than in English. There is evidence that in Italian as well, C. adopts a slot and frame approach where the copula is part of a semiproductive positional pattern where two distinct elements can be identified: an initial string containing the copula and another element (demonstrative + copula, wh-word + copula, existential *ci* + copula), and an pattern-final predicative element (determiner + noun, adjective, or locative pronoun).

8.2.4. Emergence of Tense contrasts in copula forms

In Italian like in English, virtually all copula forms are either 3 p.s. or 3 p.p. (see Table III above). From a purely morphological point of view, it is quite clear that in Italian too, C. shows semi-productive use of only a third of the copula *essere*'s paradigm. His knowledge of the copula is best characterised as piecemeal learning of a small set of forms in construction-specific contexts, rather than across-the-board

realisation of functional Agreement and Tense features. Unlike in English however, there is some emerging contrastive use of imperfect tense copula forms in the Italian data from file 13 (2;5.6) onwards, as shown in Table VII below:

File	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.
13	-	-	7	-	-	-
14	-	-	1	-	-	-
15	-	-	4	-	-	1
16	2	-	3	-	-	-
17	-	-	-	-	-	-
18	-	-	3	-	-	-
19	-	-	-	-	-	-
20	-	-	12	-	-	1

Table VII. Distribution of imperfect tense copula *essere*.

Here too, as was the case for the present tense copula, the vast majority of forms are 3 p., and specifically 3 p.s.. Out of a total of 34 imperfect tense copula *essere* forms, 30 (88.23%) are 3 p.s., 2 are 1 p.s. (6%), and 2 (6%) are 3 p.p.. Almost half of the 3 p.s. forms (1 in file 18, 2;11.12, and 12 in file 20, 3;0.17), and one of the 3 p.p. forms in file 20 are found with existential *ci* (*c'era/c'erano, "there was/there were"*) and are strictly correlated with the emergence of narratives in the past:

(13) File 20

*CAR: c'era una volta un cavallo che [/] che andava qua dentro.
%eng: there was once a horse that [/] that went inside here.

(14) File 20

*CAR: poi una volta c'era un ippopotamo.
%eng: then once there was a hippo.

(15) File 20, line

*CAR: c'erano tante [/] tante giraffe.
%eng: there were many [/] many giraffes.

Although the number of past copula forms accounts for under 7% of all copula tokens, nevertheless they are used appropriately to refer to past events, and they are used contrastively with present tense forms. The relatively high number of *c'era* forms could *prima facie* be problematic. Because they are typically used in stereotyped beginnings of story-telling (corresponding to English *"Once upon a time there was a ..."*), one might reasonably suspect that they have been memorised by C. as unanalysed

forms which do not necessarily carry a specified [+Past] Tense feature. Examples such as (13) however, clearly show that the *c'era* forms are analysed as forms carrying a [+Past] Tense feature. The imperfect *andava, "(it) went",* in the relative clause is correctly inflected for 3 p.s. Agreement features, and for [+Past] Tense feature which matches the [+Past] Tense feature in the matrix clause containing the *c'era* form. There are in fact five cases altogether in which a [+Past] copula form cooccurs with another [+Past] form in the same utterance. In the example in (16) below the past copula is in the embedded clause, its [+Past] Tense feature is selected by the present perfect form in the matrix clause:

(16) File 20

*CAR: Claudia mi ha detto che +//.
%eng: Claudia has told me that +//.
*CAR: +, non era sporcopino.
%eng: +, it wasn't sporcupine.

Despite the small number of past copula forms, there is some evidence that these forms are being appropriately contrasted in terms of [± Past] Tense features. In contrast to present Tense forms where there is an error rate of 14.55% (see Table VI above), imperfect past forms always agree in person and number with their subject. Recall however that all of the errors in present tense forms of the copula are to be found with 3 p.p. forms, where 3 p.s. forms are used instead. C.'s difficulties in number agreement are unidirectional, and specifically involve existential copula constructions. Therefore the 100% correct score on imperfect tense forms could be inflated by the very small proportion of plural forms (2/34).

In sum, the data presented in this section show that C. can use both present tense forms of the copula *essere*, and imperfect past forms, albeit the past tense forms emerge later and are much less frequent. For both present and past forms, there is no across-the-board mastery of the paradigm, it is only 3 p.s. and 3 p.p. contexts that are represented in a significant way; 2 p.s. and 2 p.p. are not attested at all, either in the present or in the past. A certain degree of construction-specificity in the use of the copula is found in the data overall; especially noteworthy is the absence of subjects realised by DPs containing a determiner and a common noun, proper names, and pronouns.

Almost half of all past tense forms are represented by the existential construction which is typically found as the formulaic beginning of fairy tale narratives

(*c'era /c'erano una volta ..., "once upon a time there was/were ..."*). Despite the formulaic origin of such copula constructions there is some evidence that C. has analysed these forms correctly in terms of their Agreement and Tense specifications, and that he can use them contrastively. Overall, the evidence analysed so far does not conclusively show that abstract Agreement and Tense features are being morphosyntactically spelled out by the use of the copula. Although the error rate is low, the restriction of copula forms largely to 3 p.s. and 3 p.p. contexts indicates a somewhat narrow and circumscribed knowledge of the paradigm. In the next section C.'s performance on present tense forms of lexical verbs will be investigated in detail to assess to what extent the notion of subject-verb Agreement is operational in the child's grammar.

8.3. The marking of Tense and Agreement in Italian lexical verbs

8.3.1. Inflectional and periphrastic verb forms

Unlike English, which only marks Number in the present tense paradigm, Italian is a Person-marking language where each of the six different person/number combinations takes a unique inflectional ending. Verbs are classified into three conjugations: the *-are* conjugation (e.g. *parl-are, "to speak"*), the *-ere* conjugation (e.g. *ved-ere, "to see"*), and the *-ire* conjugation, which includes a subgroup of verbs that take an *-isc* infix between the stem and the inflectional affix (e.g. *cap-ire, "to understand"*).

Italian verbs never appear as unmarked bare stem forms like English verbs (e.g. *talk*). Moreover, while English stems are monomorphemic, Italian stems are bimorphemic: the verbal root is followed by the thematic vowel of the relevant conjugation, as exemplified in Table VIII. The infinitival forms *parlare, vedere,* and *capire* can be segmented into the roots *parl-*, *ved-*, *cap-*; and the stems *parl-a, ved-e, cap-i,* where the vowels *a, e* and *i* indicate the verb's membership to one of the three conjugations, while the suffix *-re* marks the form as infinitival.

conjugation	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.
parl-a-re	parl-o	parl-i	parl-*a*	parl-iamo	parl-ate	parl-ano
ved-e-re	ved-o	ved-i	ved-*e*	ved-iamo	ved-ete	ved-ono
cap-i-re	cap-isc-o	cap-isc-i	cap-isc-e	cap-iamo	cap-ite	cap-isc-ono

Table VIII. Indicative present

Table VIII above, and Table IX and Table X below schematise three of the simple paradigms which will be relevant for the discussion of C.'s data.

conjugation	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.
parl-a-re	parl-avo	parl-avi	parl-ava	parl-avamo	parl-avate	parl-avano
ved-e-re	ved-evo	ved-evi	ved-eva	ved-evamo	ved-evate	ved-evano
cap-i-re	cap-ivo	cap-ivi	cap-iva	cap-ivamo	cap-ivate	cap-ivano

Table IX. Indicative imperfect

conjugation	2 p.s.	1 p.p.	2 p.p.
parl-a-re	parl-a	parl-iamo	parl-ate
ved-e-re	ved-i	ved-iamo	ved-ete
cap-i-re	cap-isc-i	cap-iamo	cap-ite

Table X. Imperative

The full inflectional system includes 21 simple and compound, finite and non-finite forms, 16 of which are commonly used. Compound or periphrastic forms require the use of auxiliary *essere* or *avere* and are followed by a past participle, when *essere* is used the past participle must agree in number and gender with the subject:[3]

(17) La ragazza è andata.
"La-s.fem. ragazza-s.fem. is gone-3-s.fem."

(18) I ragazzi sono andati.
"The-p.masc. boys-p.masc. are gone-p.masc."

If *avere* is used the past participle is in its unmarked form (masculine, singular), but it must agree with the object of a transitive verb if the object is cliticised and precedes the inflected auxiliary:

(19) Ho visto la casa.
"(I) have seen-sing.masc. the house"

(20) L' ho vista.
"It-sing.fem (I) have seen-sing.fem."

Progressive forms employ the auxiliary *stare* and an invariable gerundival participle:

(21) Laura sta parlando a Paolo.
"Laura is talking to Paolo"

[3]Auxiliary selection depends on whether the verb is ergative, in which case it will select *essere*, or whether it is unergative and will therefore select *avere* (Burzio, 1986).

Tables XI and XII list past participles and gerundival participles for the three conjugations.

conjugation	past participle
parl-a-re	parl-ato
sed-e-re	sed-uto
cap-ire	cap-ito

Table XI. Past participles

conjugation	gerundival part
parl-a-re	parl-ando
sed-e-re	sed-endo
cap-i-re	cap-endo

Table XII. Gerundival participles

8.3.2. The status of inflectional morphology in acquisition

Previous studies on the acquisition of Italian in monolingual children (Hyams, 1986, 1992; Pizzuto & Caselli, 1992, 1994; Cipriani, Chilosi, Bottari & Pfanner, 1993; Caselli, Leonard, Volterra & Campagnoli, 1993; Guasti, 1993/94; Antelmi, 1997) have reported contrasting findings on the mastery of verbal morphology. On the one hand, studies like Hyams's (1986) and Guasti's (1993/94) conclude that Italian-speaking children master inflectional verbal morphology very early and consistently:

> "In Italian the verb is inflected to agree in person and number with the subject ... With rare exceptions the verbs in early Italian are correctly inflected.... agreement is not restricted to a particular grammatical person ... we find examples of each grammatical person, with the exception of the 2nd person plural" (Hyams, 1986: 135)

> "The Italian data prove that Italian children not only distinguish between finite and infinitive verbs, ... but they also know the agreement system... More precisely, Italian children use the inflection for person very early, especially with the singular persons. With few exceptions, these inflections are correctly distributed among subjects." (Guasti, 1993/94: 21)

Nonetheless, a number of researchers have challenged the conclusions reached by Hyams, and therefore, albeit indirectly, those drawn independently by Guasti. Pizzuto & Caselli (1992,1994), and Caselli, Leonard, Volterra & Campagnoli (1993) characterise the acquisition of Italian morphology as a gradual process of piecemeal learning. Pizzuto & Caselli (1992) use longitudinal data from three monolingual Italian children (age range 1;4-3;0) to assess to what extent the morphological paradigms of verbs, pronouns and articles can be said to be acquired by the children, and, if so, at what age. In their analysis they distinguish between "first appearance" and "point of

acquisition" for a given morpheme which, following Cazden (1968), is defined as the first of three consecutive samples where the morpheme is produced correctly in at least 90% of obligatory contexts. Two additional criteria are also used in determining point of acquisition: firstly, each of the samples that are crucial for scoring the acquisition point must contain at least five obligatory contexts of use; secondly, with specific reference to verbal inflections, a person inflection is considered to be used productively if it appears with at least two different verbs in each of the samples chosen for the scoring of the point of acquisition. Given these stringent productivity criteria, Pizzuto & Caselli observe that none of the major inflectional paradigms under investigation is fully mastered by the children in their study. As regards the production of verb inflections, only a small set of singular present tense inflections can be said to be acquired and be used productively: 1 p.s. and 3 p.s. present indicative, 2 p.s. present indicative/imperative, 3 p.s. present indicative copula for two of the children, and only 3 p.s. present indicative for the third child in their study.

The major conclusion that Pizzuto & Caselli draw from the children's incomplete mastery of the verbal inflectional paradigms is that they cannot be credited with productive, across-the-board knowledge of subject-verb agreement as claimed by Hyams (1986). They are also especially concerned with the total lack of productivity of plural person inflections, to them this constitutes particularly strong evidence against a general rule of subject-verb agreement operating in the children's grammars. In a reply to the Pizzuto & Caselli's article, Hyams (1992b) dismisses the theoretical relevance of the absence of plural verbal inflections by arguing that "all we can conclude from the lack of plural inflections is that children do not like to talk about plural things" (Hyams, 1992b: 698).

Similar arguments against across-the-board acquisition of verbal inflections have also been put forward by Pine & Lieven (1993, 1997), Lieven, Pine & Baldwin (1997), Pine, Lieven & Rowland (1998), and Tomasello (1992) for English, Rubino & Pine (1998) for Brazilian Portuguese, and Gathercole, Sebastián & Soto (1999, in press) for Spanish. The work on English by Lieven, Pine and colleagues goes back to some ideas originally put forward by Braine (1976) on children's early multiword utterances as limited scope formulas that specify the relative ordering of narrowly circumscribed semantic classes. Limited scope formulas are best characterised by the appearance of a constant, such as *big* or *see*, and an open position that can be filled by a large number of lexical items. Capitalising on Braine's "frame + slot" approach in a study of the first 400 multiword utterances of 11 monolingual English children

between the ages of 1;0 and 3;0, Lieven, Pine & Baldwin (1997) claim that a positional analysis can account for a mean of 60% of all the children's multiword utterances, and that the vast majority of all other utterances are defined as frozen forms.[4] Lieven et al.'s study focuses on pronoun case-marking, and verb-argument structures in two-place predicates. In both cases their conclusion is that children operate with distributional patterns built around specific lexical items. There is no evidence that they have systematic control over Case-marked pronouns as would be reflected by underlying syntactic knowledge of case checking requirements. As for verbs, they note again that the basis of organisation of the children's vocabulary revolves around specific verbs rather than relying on systematic semantic and/or syntactic regularities.

Lieven et al.'s conclusions are along the same lines as those formulated by Tomasello (1992) in an in-depth study of an English-speaking child's acquisition of verbs. The analysis of the child's acquisition of verb morphology and argument structure reveals a significant degree of lexical specificity. It is as if individual verbs were independent islands of organization where any morphological and syntactic information is initially applied on an item-by-item basis and no generalised productive processes are at work. The Verb Island Hypothesis (Tomasello, 1992) argues for the verb-specific nature of initial morpho-syntactic knowledge, and against category-general operations.

---

[4]Pine & Lieven (1993), and Lieven, Pine & Baldwin (1997) rely on the positional regularity (P) of particular lexical items (the "frame"), and evidence that a particular lexical item appears independently in the child's vocabulary rather than embedded in a rote-learned phrase (i.e. in a "slot"). The coding scheme adopted to identify frozen, intermediate, and constructed utterances is the following:

a) a multiword utterance is defined as "frozen" if there is no evidence that its lexical items also appear independently elsewhere.

b) "intermediate" utterances are of two types:

(i) they are the second instance of positional regularity for one of the words (P + [2]), and the other word(s) have already occurred independently in the child's lexicon (I+), or

(ii) all of the words have already occurred independently elsewhere (I+;I+)

c) "constructed" utterances include those multiword utterances where they are the third instance of positional regularity for one of the items, and there is evidence for the independent occurrence of the other word(s) elsewhere (I+; P+ [3]). Once a pattern is defined as constructed it constitutes a "frame" consisting of a lexical item in a fixed position with a slot on one or other side of it.

Even for languages with a richer array of inflectional affixes than in English, such as Italian, Spanish and Catalan, for which there seems to be evidence that children acquire verbal morphology very early, there have recently been proposals for piecemeal learning along the lines of that proposed by Tomasello, and Pine & Lieven and colleagues for English. Evidence for the lack of early, across-the-board productive use of verbal morphology in Italian has been provided by Pizzuto & Caselli (1992, 1993) and Caselli, Volterrra, Leonard & Campagnoli (1993), as mentioned above.

Recent work on the acquisition of Spanish in two monolingual children by Gathercole, Sebastián & Soto (1999, in press) has also shown that there is little productive command of verbal morphology in Spanish ECG and that children begin with a single form per verb. Gathercole et al. (1999) adopt two measures of productivity that have been adopted previously by Pizzuto & Caselli (1994) and by Fernández-Martinez (1994): the same verb root must appear in at least two inflected forms, and the same inflection must appear with at least two verb types. When elements of the verb paradigm do start to be used productively, they do so in a very gradual verb-specific fashion. Although the two children in the Gathercole et al.'s study use a relatively large number of verbs in their speech as early as 1;6, nevertheless when productivity criteria are applied one child does not display any contrastive command of any inflection until 1;11, and the other until 2;1. Although both children start out with one verb form per verb type, there is however a gradual increase in the proportion of verbs that appear in more than one form. As for the kind of contrast of Person, Number, and Tense features the authors state that "it is impossible to say that one of these emerges, is learned, or is filled in before the others. There is no across-the-board acquisition of any of these" (Gathercole, Sebastián & Soto, 1999: 17).[5]

Although both of these Spanish-speaking children show considerable gaps in the mastery of verbal paradigms, the possibility still remains that they do not produce

[5]In a study of four monolingual Catalan-speaking children and one monolingual Spanish-speaking child, Grinstead (in press) argues that the Tense and Number Phrases do not form an active part of the clause structure of Child Catalan and Child Spanish in the earliest stages. The Person Phrase, by contrast, does appear to be active from the beginning of acquisition. This argument for the early realisation of a Person Phrase is based on the observation that while the children in the study produce a number of verbs inflected for 1 p.s., 2 p.s., and 3 p.s., plural forms are initally very rare. What is missing from Grinstead's data analysis however, is evidence that verbs that appear with one Person inflection, also appear with other Person inflections at the same point in time. The fact that, for example, the 1 p.s. inflection appears with a number of different verb types is not in itself a guarantee that 1 p.s. is used contrastively. It is possible, for exampel, to have a situation in which there is no overlap between the verbs that appear with 1 p.s. and the verbs that appear with other inflections. If this is the case, it is obviously inappropriate to speak of mastery of Person contrasts, and of the realisation of a Person Phrase.

certain forms because they have no occasion to do so, or simply because they deliberately avoid them, as claimed by Hyams (1992). In order to rule out this possibility, Gathercole et al. have searched the data for commission errors that would show that the children are not simply avoiding certain forms, for whatever reason, but that they have not actually learned them yet. They show that for both children the number of commission errors is higher in the sessions prior to the acquisition of contrastive use of verb forms, and decreases thereafter. Thus the lack of productive command of an inflection is not simply due to the absence of that particular inflections. In fact it is the case that commission errors do occur, and there are indeed cases in which the wrong inflection is produced for a given obligatory context. It must be noted however that for one of the two children the figures are so low that their statistical significance may be called into question. Maher produces a small number of verb tokens overall and also a small number of errors: 3/6 (50%) prior to the "contrast" session and 3/7 (42.9%) after the "contrast" session. By contrast, Maria's error rate decreases from 12/34 (35.3%) prior to the "contrast" session. to 12/145 (8.3%) after the "contrast" session, a somewhat more meaningful result.

These relatively high overall error rates contrast sharply with other findings in the literature on the acquisition of other morphologically rich languages such as Italian and Brazilian Portuguese. For Italian, Pizzuto & Caselli (1992) have found overall error rates oscillating from 1 to 4% in the three children in their study (depending on the child and the verb class). Rubino & Pine (1998) have also found a fairly low overall error rate in the use of inflected verbal forms in the speech of one child acquiring Brazilian Portuguese (3%). Because the samples in these two studies are limited to one child (Rubino & Pine, 1998), at most three children (Pizzuto & Caselli, 1992), it is clearly difficult to disentagle the normative significance of these results from potential biases due to individual variation. Even in Gathercole et al.'s study, where the same methodological tools have beeen applied to the speech samples of two children, we find markedly different results as regards overall error rates, a situation which is also potentially biased by the low overall number of both correct and incorrect tokens for Maher.

In the following sections, the same methodological tools described above which have already been successfully applied to the study of the acquisition of verbal inflections in Italian and Spanish, will be applied to the analysis of C.'s verbs.

8.3.3. Inflected verbs forms in C.'s Italian

In line with previous studies on the acquisition of morphology in Romance languages, the relative proportions of correct and incorrect tokens of inflected verb forms for the present indicative in Table XIII below reveal low error rates:

File	1 p.s.	2 p.s.	3 p.s.	1 p.p.	2 p.p.	3 p.p.	total %errors
1	-	-	-	-	-	-	-
2	-	-	1/0	-	-	-	0
3	4/1	-	7/0	-	-	-	8.33
4	3/0	-	3/0	-	-	-	0
5	-	-	-	-	-	-	-
6	1/0	-	4/0	-	-	-	0
7	2/0	-	4/0	1/0	-	-	0
8	-	-	2/0	-	-	-	0
9	3/1	1/0	4/0	3/0	-	-	9.09
10	1/0	1/0	9/1	1/0	-	0/1	14.29
11	1/0	-	5/2	-	-	0/1	25
12	11/2	-	22/0	-	-	-	6.06
13	28/0	1/0	33/0	-	-	3/1	1.53
14	12/0	9/1	33/0	5/0	-	1/0	1.6
15	315/1	5/0	24/0	8/0	-	-	1.92
16	15/1	3/0	1/0	3/0	-	-	4.54
17	11/0	4/0	8/1	-	-	5/0	3.57
18	12/0	9/0	22/0	-	-	2/0	0
19	15/0	7/0	23/0	7/0	-	15/0	0
20	7/3	-	-	1/0	-	10/1	19.1
total N tokens	141/9	40/1	172/5	29/0	-	38/5	
error rates	6.38	2.5	2.9	0	-	13.88	

Table XIII. Proportion of correct/incorrect verb tokens in the present indicative

The overall error rate for C.'s present indicative forms is 5.13%, oscillating between 0% for 1 p.p. forms to 13.88% for 3 p.p. This figure is in line with the overall error rates previously found in other studies such as Pizzuto & Caselli's (1992), and Rubino & Pine's (1998). Over time, the mean error rate across the various sessions ranges from 0% to a peak of 25% in file 11, where however the combined number of incorrect and correct tokens is small to have any major statistical significance, only 8 incorrect verbs altogether.

Contrary to what may at first be expected, C. is very accurate from the earliest recordings, out of a total of 7 sessions with 0% error rate, 5 are found before 2;2.3. It is nevertheless important to emphasise once again that low error rates without any indication of actual productivity cannot be taken as good evidence of real mastery.[6] It

[6] Antelmi (1997:143) reports errorless verb production in the early stages (1;6-1;8) of acquisition for Camilla , a child acquiring Italian monolingually, followed by a very high agreement error rated between 1;8 and 1;11 (43% to 51% of incorrectly agreeing forms).

could simply be that C. is repeating a small number of rote-learned forms which do look adult-like, but which are not indicators of true competence. A more fine-grained analysis of the data is in order to find out to when and to what extent C. actually displays any productive control of inflected forms. The same criteria of productivity previously adopted in the literature will be used here: in order for an inflection to be used productively it must be correctly used at least twice with at least two different verb types in the same session.

Table XIV shows the number of tokens, types and verb forms across the 20 Italian files broken down by inflection. The cells include the number of tokens for a given inflection and below the number of verb types that are found with that inflection, e.g. in file 6 there are 4 verbs inflected for 3 p.s. indicative present, and 3 different verb types: 2 tokens of *mangia, " (s/he)eats"*, 1 token of *vola, "(s/he/it) flies"*, and one token of *porta "(s/he) brings"*. The rightmost column indicates the total number of verb tokens for each session, followed by the total number of different verb types, and the number of different verb type forms. For example in file 10 we have a total number of 17 verb tokens, 8 verb types, and 9 verb type forms; one verb type (volere, "to want") appears with 3 different verb forms: *volo (= voglio), "(I) want"* (1 p.s. Pres. Ind.), *vuoi, "(you) want"*, (2 p.s. Pres. Ind.), and *vole (= vuole) "(s/he) wants"*, (3 p.s. Pres. Ind.).

File	Inf.	Bare P.P.	Imperative sing.		Present						Present Progressive						Imperfect						Present Perfect						Verb tokens/ Verb types/ Verb forms
			Aff	Neg	sing. 1	2	3	pl. 1	2	3	sing. 1	2	3	pl. 1	2	3	sing. 1	2	3	pl. 1	2	3	sing. 1	2	3	pl. 1	2	3	
1																													
2			1			2																							3/3/3
						2																							
3					4	7																							11/5/6
					2	4																							
4	3		39		3	3																							48/5/6
	1		2		1	2																							
5			7																										7/2/2
			2																										
6			8		1	4																							13/6/6
			2		1	3																							
7		4	2		2	4	1																						13/8/8
		2	1		2	2	1																						
8			1			2																							3/3/3
			1			2																							
9			2		3	1	4	3																					13/8/8
			2		2	1	1	2																					
10			5		1	1	9	1																					17/8/10
			1		1	1	6	1																					
11			9		1		5																1						16/6/7
			1		1		4																1						
12					11		22																	2					35/10/11
					4		5																	2					
13	4	4	4		28	1	33			3		8		5										6				7	103/22/33
	2	3	1		8	1	6			3		5		2										3				1	
14		2	3		12	9	33	5		1		6											5		15	1	2		93/25/38
		1	1		5	6	6	3		1		5											4		4	1	1		
15	19		7		15	5	24	8				4											5	1	13	1			102/23/45
	8		3		4	3	8	2				3											3	1	6	1			
16	5				15	3	1	3									1						9		4	2			43/15/21
	2				5	3	1	3									1						4		3	1			
17	9		1		11	4	8			5		4					5						4	1	4	1			47/22/31
	5		1		3	2	5			2		4					1						4	1	2	1			
18	7	2	1		12	9	22			2													5		7			2	69/23/34
	3	2	1		9	4	8			1													2		4			2	
19	17			6	15	7	23	7		15													4	5	3	2			104/24/45
	7			2	6	5	5	4		5													3	3	3	2			
20	18	2	6	4	11	12	11	11		12							1	19		3			1		10	5			126/36/66
	11	2	2	2	5	4	6	2		3							1	9		2			1		7	3			

Table XIV. Distribution of Italian verb tokens

At first glance, the data in Table XIV suggest that C. is using a wide range of inflections with a variety of verb types. This being the case, one might be tempted to credit the child with sophisticated knowledge of subject-verb agreement, a morphological reflex of the presence of an AGR category in his grammar to which inflected verbs must raise to check their phi-features. Guasti (1993/94) makes exactly such a claim when reviewing the performance on subject-verb agreement of three Italian-speaking children between the ages of 1;8 to 2;7. Just like in C.'s sample, the children studied by Guasti use a considerable number of inflected verbs in the present indicative from the earliest recordings, and the overall error rate ranges from 1% to 3% across the children; finally, the children do use person inflections with a number of verb types, an indication that they have not simply rote-learned the forms that they use correctly. Guasti concludes that "[t]hese observations, combined with the finding that agreement inflections are used correctly, support the claim that the grammatical notion of subject-verb agreement and the ability to make a morphological analysis of verbal forms are in place at around the age of 1;11 to 2;1" (Guasti, 1993/94: 29-30).

Although Guasti's conclusions are not unreasonable in the light of the data she presents, nevertheless her only two criteria for establishing productive use of various inflections are: a) low error rates, and b) the anecdotal evidence that the children use inflected verbs with a variety of verb types. The reliance on low error rates and the absence of a sound quantitative analysis are problematic for a number of reasons. Firstly, it is impossible to assess to what extent the children in the study actually do consistently use a number of different verb types for any given person inflection. Guasti only gives a handful of examples which may not be representative of the child's systematic behaviour. Secondly, and most importantly, this productivity criterion, i.e. the observation that a given inflection appears with a number of different verb types, is not by itself adequate to verify whether children can also use the forms in a contrastive way. Suppose that a child used the 3 p.s. indicative present inflection in a given session with 5 different verb types, for example *mangia, "s/he eats", dorme "s/he sleeps", corre, "s/he runs", salta, "s/he jumps", guarda "s/he looks"*. For Guasti this would be clear evidence that the child is using the 3 p.s. indicative present inflection productively; there are no errors and the inflection appears with a variety of different verb types. However, as Tomasello (1992) has observed, it is of paramount importance to find *contrastive* uses of the same verb type in order to decide to what extent the morphology and the argument structure that go together with that verb are really productive:

> "In what sense is *working* a present progressive if T. never once used *work*? There are other problems as well, but for now let me make clear that I am attempting to establish ... whether and at what time T. actually contrasted two forms of the same 'verb'. This requires that both forms be used during roughly the same time period and that they be used in similar conceptual situations" (Tomasello, 1992: 161)

Similar arguments are provided by Pizzuto & Caselli (1992, 1994), and by Gathercole, Sebastián & Soto (1999, in press) in their studies of the acquisition of Italian and Spanish verbal morphology. Both the fact that a given inflectional morpheme appears with different verb types, *and* the fact that the same verb appears with at least two different inflections in the same session, are essential to establish a reliable measure of productive use. Taken together these two criteria offer a more explicit definition of acquisition than simply the 90% correct forms in obligatory contexts which is all too often used as the only method to assess productive use of morphology. As has already been shown by previous studies on the acquisition of Italian and Spanish, if these productivity criteria are applied to child data, it is not so clear that children learning morphologically rich languages acquire verbal paradigms in an across-the-board fashion much earlier than children acquiring less richly inflected languages like English. More importantly, it seems that even Italian- and Spanish-speaking children go through a stage where subject-verb agreement is not errorless, and where the acquisition of Person, Number and Tense inflections takes place on a verb-specific basis. On the basis of these empirical facts there is reason to question the assumption that in more richly inflected languages an Agreement projection is justified from the earliest verbal utterances.[7] In fact, what these studies have in common is the conclusion that children learn morphology in a piecemeal fashion, where inflections are being learned on a verb-by-verb basis, and where inflected forms are often not used contrastively by the child.

With these clarifications in mind we can now proceed to examine C.'s data in the attempt to establish whether and when he has acquired productive and contrastive knowledge of particular inflections. The crucial issue is whether his competence is best

---

[7]Of course the correspondence between the overt morphological realisation of subject-verb agreement on verbs, and the presence of a functional projection where such overt morphological features can be checked, rests on the assumption that the postulation of FCs in a child's grammar is justified only as long as there is empirical morphosyntactic evidence to support it (see Grimshaw, 1994). This argument is at the core of the WC approach advocated in this work, but we are aware that proponents of other models of language acquisition, such as SC approaches, question this direct relationship between morphophonological learning and the realisation of abstract syntactic notions. Their argument is that FCs are a necessary part of the child's grammar, the morphological realisation of FCs is simply the result of lexical learning, a process that in itself might take some time to accomplish. As will be pointed out in chapter 9, this dissociation between grammar and the lexicon is a spurious one, and it is undesirable for both empirical and theoretical reasons.

characterised in terms of piecemeal lexical-specific learning, or whether there is evidence that the child shows mastery across the verbal paradigms.

8.3.4. Verb-specific and paradigm-based acquisition of inflections

As shown in Table XIV, until file 10 (2;2.17) there is virtually one verb form per verb type per session. Although file 3 (1;11.17) and 4 (1;11.25) look like two exceptions, there is reason to dismiss them as invalid examples of productive and contrastive use of verbal inflections. In file 3 one verb type, *volere (to want)*, is found with both a 1 p.s. present indicative inflection, *volo (= voglio) ((I) want)*, and a 3 p.s. present indicative inflection, *vole (= vuole) ((s/he) wants)*, this would normally indicate contrastive use, but not in this case since the 3 p.s. present indicative form *vole* is erroneously produced instead of a 1 p.s. present indicative form, since this is an error it is automatically excluded from the count of productivity. As for file 4 the verb *prendere* appears both in the infinitival form, and with a 1 p.s. present indicative inflection. Again the three tokens of the infinitive must be eliminated from the count of productive forms because they are instances of Optional Infinitives, and as such an ungrammatical form (see sections 4.3.2. and 4.3.3.).

Starting from file 10 (2;2.17), there is an observable tendency for a number of verb types to appear with more than one form, and this trend becomes particularly clear from file 13 (2;5.6) onwards. Because the rightmost column in Table XIV only lists the total number of types and the total number of forms, when there are more forms than types it is not clear whether it is the case that one verb type appears with a large number of forms, or whether there are several verb types which minimally appear with two different inflections. This is a subtle but interesting difference in that it would reveal two different acquisitional strategies. In the case in which the larger number of verb forms were due to one, or two, verb types appearing with multiple forms it would look as if the child were filling in a paradigm on a verb-by-verb basis. By contrast, if one found that a large number of verbs appear with only two or at most three different forms, it would suggest that the child is using the inflections in a more across-the-board, paradigmatic fashion. In the former case it is as if the child were filling in the slots of the paradigm for a particular verb ; in the latter it is as if the child were filling in the verbal paradigm itself, with less regard to the build-up of a repertoire of inflected forms for specific verbs. In order to disambiguate the two possibilities a complete list of all the verb types is provided in Appendix 3.

Appendix 3 includes the 82 verb types that appear in C.'s Italian corpus listed in alphabetical order. For each verb the table indicates the verb type, e.g. *aiutare, "to help"*, and the number of verb form(s) that are present in each file for that particular verb. For instance *aiutare* appears with two different verb forms, 2 p.s. present indicative and 1 p.s. present indicative. There is one token of 1 p.s. present indicative in file 18, one token of 1 p.s. present indicative in file 19, and one token of 2 p.s. present indicative in file 19. The last row in the table in Appendix 3 also tabulates the number of verb types for each session, the number of verb types that are used contrastively, and the percentage of verbs that are used contrastively. Disregarding file 3 and 4 where the contrasting forms are in fact errors, in file 10, 11, and 12 we find the first occurrences of a single verb type used with different forms. From file 13 (2;5.6) onwards there is a consistent proportion of verbs that appear with more than one form, ranging from 20% to a peak of 56% in file 15 (2;9.6). Although there is a fairly abrupt transition around 2;5.6, the progression is not a linear one and the proportion of verbs used contrastively varies quite considerably from one session to the next. A closer inspection of the distribution of multiple inflections across verb types reveals that 17 verb types out of the total number of 82 (20.73 %), occurr with a total of four or more forms across the 20 files as shown in Table XV:

V type	N of V forms
FARE to do/to make	13
ANDARE to go	9
PRENDERE to take	8
METTERE to put	7
GIOCARE to play	7
MANGIARE to eat	6
AVERE to have	6
DIRE to say	6
DARE to give	5
DOVERE must	5
POTERE can	5
VOLERE to want	5
CADERE to fall	4
CHIUDERE to close	4
GIRARE to turn	4
GUARDARE to look	4
STARE to stay	4

Table XV. Verb types with four or more forms

Looking at the data from a longitudinal perspective it does seem as if C. is working around specific lexical items by adding new forms to fill in the paradigm for that particular verb. It is particularly noticeable that the verb *fare (to do/to make),* makes the largest contribution to verb contrasts across the various files. It is the only verb that is consistently found in all files from 13 onwards with a number of contrastive inflected forms, ranging from 4 to 7, which are all also productively used with other verb types.

8.3.5. The emergence of grammatical contrasts

Table XVI shows the number of verb types for each productive verb form for which there is also contrastive use.

File	Inf. Bare P.P.	Imperative sing. Aff Neg	Present sing. 1 2 3 pl. 1 2 3	Present Progressive sing. 1 2 3 pl. 1 2 3	Imperfect sing. 1 2 3 pl. 1 2 3	Present Perfect sing. 1 2 3 pl. 1 2 3	Verb tokens/ Verb types/ Verb forms
1							
2							3/3/3
3							11/5/6
4							48/5/6
5							7/2/2
6							13/6/6
7							13/8/8
8							3/3/3
9							13/8/8
10			1				17/8/10
11			1				16/6/7
12			1				35/10/11
13	2		3 2	2 1		2	103/22/33
14			1 4 4 3	2		3	93/25/38
15	5	2	3 3 8 2	1 1		5	102/23/45
16	1		3 1 3			1 1	43/15/21
17	4		1 2 1	1		3	47/22/31
18	1		4 3 4			1 4 1	69/23/34
19	5		5 4 2 3 6			2 2 2 1	104/24/45
20	6		2 4 2 1 2		6 1	4 1	126/36/66

Table XVI. Productive Italian verb types

Although in file 10 (2;2.17), 11 (2;3.7), and 12 (2;4.14) the 3 p.s. present indicative inflection is used productively, it is not until file 13 (2;5.6) that the first contrasts emerge. A number of contrasts make their appearance simultaneously: there is a Mood contrast between infinitive and indicative, a Person contrast between 3 p.s. and 3 p.p. in the present indicative, an Aspect contrast between present indicative and the present progressive, another Person contrast in the present progressive between 3 p.s. and 3 p.p., and finally a Tense/Aspect contrast between 3 p.s. present indicative and 3 p.s. present perfect. Table XVII schematises the appearance of Person, Number, Tense, Aspect and Mood contrasts over time

File	Person	Number	Tense	Aspect	Mood
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13		**3 p.s./3 p.p.Pres. 3 p.s./3 p.p. Pres. Prog.**	**Pres./Pres. Perf.**	**Pres./Pres. Prog./Pres. Perf.**	**Inf./Ind.**
14	**1 p.s./2 p.s./3 p.s. Pres.**	**1 p.s./1 p.p.**	Pres./Pres. Perf.	Pres./Pres. Prog./Pres. Perf.	
15	1 p.s./2 p.s./3 p.s. Pres. 1 p.s./3 p.s. Pres. Prog.	1 p.s./1 p.p. Pres.	Pres./Pres. Perf.	Pres./Pres. Prog./Pres, Perf.	Inf./Ind./**Imp.**
16	1 p.s./2 p.s. Pres.	1 p.s./1 p.p. Pres.	Pres./ Pres. Perf.	Pres./Pres. Perf.	Inf./Ind.
17	1 p.s./3 p.s. Pres.	3 p.s./3 p.p. Pres.	Pres./Pres. Perf.	Pres./Pres./Pres. Perf.	Inf./Ind.
18	1 p.s./2 p.s./3 p.s. Pres. **1 p.s./3 p.s. Pres. Perf.**	**3 p.s./3 p.p. Pres. Perf.**	Pres./Pres. Perf.	Pres./Pres. Perf.	Inf./Ind.
19	1 p.s./2 p.s./3 p.s. Pres. **1 p.p./3p.p. Pres.** 1 p.s./**2 p.s./**3 p.s. Pres. Perfect	1 p.s./1 p.p. Pres. 3 p.s./3 p.p. Pres. **1 p.s./1 p.p. Perf.**	Pres./ Pres. Perf.	Pres./Pres. Perf.	Inf./Ind.
20	1 p.s./2 p.s. Pres. 1 p.s./3 p.s. Pres. Perf.	3 p.s./3 p.p. Pres.	Pres./Pres. Perf./**Imperf.**	**Pres.Perf./ Imperf.**	Inf./Ind.

Table XVII. The emergence of Person, Tense, Aspect and Mood contrasts

What is striking about the data in Table XVII is that, although Tense/Aspect contrasts start to become productive at the same time as Person contrasts in the Present between 3 p.s. and 3 p.p., they are limited to 3 p.s. until file 16 (2;10.18) , when a contrast between 1 p.p. Present , and 1 p.p. Present Perfect appears for the first time. The largest number of contrasts that are added over time are Person contrasts, mainly in the Present Tense, but also in the Present Perfect, and to a lesser extent in the Present Progressive. It must be noted that there is however an important qualitative difference between the acquisition of Person inflections in the present tense paradigm, and the acquisition Person contrasts in periphrastic forms. All that is required there is to learn the inflectional paradigm of an auxiliary verb, either *essere, avere* or *stare.* Although one might be tempted to credit C. with the mastery of Person contrast, and hence awareness of this Agreement feature, there is no evidence that it is relevant across the board, apart from present tense and present perfect tense. In the present progressive, Person contrasts are limited to 1 p.s, 3 p.s. and 3 p.p., and in the imperfect the only productive Person inflection is 3 p.s. Moreover, only 37% of the 53 verb types that occur in the present indicative, are found with two or more inflections throughout the period of observation, 63% only ever appear with only one Person inflection.

As for Tense and Aspect contrasts, they start to emerge together with the first Person/Number (3 p.s. vs. 3 p.p.) contrasts in the simple present and in the present progressive. For the present progressive, only 1 p.s., 3 p.s., and 3 p.p. are productively contrasted with simple present and present perfect. While for the present perfect, 1 p.s., 2 p.s., 3 p.s., 1 p.p., and 3 p.p. all contrast with the corresponding persons in the simple present by file 18 (2;11.12). Because the Italian present perfect combines Tense and Aspect morphology, the appearance of a productive form in the present perfect has been considered to contrast productively both in terms of Aspect and of Tense as indicated in Table XVII above. However, the issue of whether emerging past tense forms which combine both Tense and Aspect morphology can be considered as past markers in child language has long been debated in the literature. Bronckart & Sinclair's (1973) seminal experimental study on the production of inflectional morphology in 74 French-speaking children aged between 2;11 and 8;7, revealed a semantic bias according to which children preferred to use the present (présent) with inherently durative events, and perfective past forms (passé composé) with events that denoted clear end results. Antinucci & Miller (1976) showed a similar tendency in an investigation of the spontaneous speech of one English- and seven Italian-speaking children, whereby they concluded that "[t]he meaning of the child's past tense is at this point rather limited. He is able to encode a past event, but only if it results in a present

state.[8] Looking at this fact from a linguistic point of view, we could say that *the past 'tense' has more of an aspectual than a temporal value*" (Antinucci & Miller, 1976: 183, my italics). In a longitudinal study of four monolingual English-speaking children (from Brown's (1973) Stage I to Stage III), Bloom, Lifter & Hafitz (1980) also report that the emergence of verbal inflections is not a process that applies across the verb category, but rather it is influenced by the inherent lexical aspect of the verb.[9] Verbs denoting durative, non-telic events, such as *play, hold, ride, write* appear almost exclusively with *-ing*, the telic, durative verbs occur with *-s* (e.g. *go, hurt, need*) and the telic, punctual verbs occur with *-ed* or irregular past morphology (e.g. *get, buy, draw, find, come*).

---

[8]While Tense expresses temporal deixis, and as such locates a situation in relation to some other time (such as speech time), Aspect is nondeictic, and it characterises "different ways of viewing the internal temporal constituency of a situation" (Comrie, 1976: 3). Grammatical aspect, also known as viewpoint aspect (Smith, 1983) refers to aspectual distinction encoded by grammatical linguistic devices, such as auxiliaries and inflections. Inherent lexical aspect, also called situation aspect (Smith, 1983), refers to lexical properties of duration and telicity which are inherent on the verb. For example a verb like *love* is inherently stative and atelic, while *die* is inherently telic and punctual. Vendler (1967) proposed a four-way classification of the inherent semantics of verbs and their lexical aspectual properties:

(i) Achievement verbs: they denote events that take place instantaneously, and are inherently punctual (e.g. *recognize, die, reach the summit*, etc.)

(ii) Accomplishment verbs: they denote events which have a duration in time, but a single clear endpoint (e.g. *run a mile, build a house, paint a picture*, etc.)

(iii) Activity verbs: they denote events which have a duration and no obvious endpoint (e.g. *play, sing, dance, work, walk*, etc.)

(iv) Stative verbs: they denote states, non dynamic situations by definitions (e.g. *live, love, like, hate*, etc.)

Each of the four categories above can be defined as a combination of three semantic features: telic, punctual and dynamic:

	ACCOMPLISHMENT	ACHIEVEMENT	ACTIVITY	STATE
PUNCTUAL	-	+	-	-
TELIC	+	+	-	-
DYNAMIC	+	+	+	-

[9]Brown's (1973: 56) five grammatical developmental stages are defined by Mean Length of Utterance (MLU):

(i)

Stage	MLU
I	1.75
II	2.25
III	2.75
IV	3.50
V	4.00

A number of other studies have reported that children initially use past (or perfective) marking predominantly with telic verbs (achievement and accomplishment verbs), gradually extending its use to activity verbs (-punctual, -telic, +dynamic), and finally stative verbs (-punctual, -telic, -dynamic) (see Rispoli & Bloom, 1985; Bloom & Harner, 1989; Clark, 1996: Pine, Lieven & Rowland, 1998) . In languages, such as English and Italian, that have progressive aspect children begin by using progressive morphology with activity verbs, and then they extend it to accomplishment and achievement verbs, but never to stative verbs. These facts are consistent with the hypothesis that children's first use of past morphology is severely conditioned by the inherent lexical aspect of verbs, and that rather than marking Tense distinctions, children are really marking Aspect or the end state resulting from that action (the Aspect before Tense Hypothesis).

More recently, Andersen (1993) and Shirai & Andersen (1995) have proposed a less stringent version of the Aspect before Tense Hypothesis, which they simply call the Aspect Hypothesis. They propose an explanation of children's first use of past tense morphology that relies on a prototype account and on the Distributional Bias Hypothesis. Shirai & Andersen's (1995: 758) prototype account for the acquisition of past morphology rests on the claim that "children acquire a linguistic category starting with the prototype of that category, and later expand its application to less prototypical cases". The three children in their study started out by using past morphology with [+punctual], [+telic], and [+result] verbs, and at the same time they used progressive morphology with prototypical activity verbs, and only gradually did they extend use of both inflections to other verbs. Shirai & Andersen also looked at the distribution of past and progressive inflections in the child-directed speech of the adult interlocutors, and there too they found a correlation between past inflection and accomplishment and achievement verbs on the one hand (58-64% of all instances of past inflection are found with accomplishments and achievement verbs), and between progressive inflection and activity verbs on the other (53-61% of progressive forms are with activity verbs). Given the distributional of verb types with past and progressive forms in the adult input, one cannot rule out that children's initial verb-specific use of past and progressive inflections does not reflect, at least partially, what they are exposed to. Shirai & Andersen (1995: 759) conclude that children's early use of past morphology encodes either "underextended past tense; [or] aspectual features such as completive, perfective, telic, punctual and so forth". They are unwilling to claim that it is either Tense or Aspect that children express when they use past morphology, the prototype of the aspectual category perfective is after all very similar to that of the category Past. It

is reasonable therefore to assume that what children are doing, at the beginning, is indeed marking Tense, albeit with the limitations of a narrower lexical category than in the adult grammar. While in the adult grammar, at least in principle, any verb can take past morphology, in the child grammar it is a subset of the adult verb category that can take past morphology, and specifically a category of verbs that has the features [+telic], [+punctual], [+result]. Once again, we see how the expression of a grammatical contrast on verbs is heavily constrained by the acquisition of specific lexical items, and not by the abstract realisation of a verb category.

In C.'s use of past morphology in Italian we see a developmental trend which is similar to that observed for his English-speaking peers. Out of 30 different verb types that occur in the present perfect, 70% are achievement verbs (N = 21) (e.g. *trovare, "to find", nascondersi, "to hide", perdere, "to lose"*), 13 % are accomplishment verbs (N = 4) (e.g. *fare, "to make", finire, "to finish"*), and the remaining 17% (N = 5) are activity verbs (*ballare, "to dance", giocare, "to play", leggere, "to read"*, and *mangiare "to eat"*)[10]. The figures clearly indicate a strong distributional bias towards the appearance of telic verbs with present perfect morphology. In particular, the large proportion of achievement verbs suggest that the [+punctual] feature may well play a major role in the use of present perfect forms. In C.'s Italian data too, there is evidence for a prototypical bias and a distributional bias in the use of past morphology, whereby prototypically [+telic] achievements verbs, and to a lesser extent accomplishment verbs, are the only verb types that are found in the present perfect.

Unlike in English, where Tense and Aspect distinctions are conflated in past inflectional morphology (e.g. *she went home*), in Italian, Tense and Aspect marking are encoded by different morphosyntactic markers in present perfect forms.[11] In a present perfect form such as *E' andata a casa, "She went home"*, the past participle encodes the aspectual information, while the tensed auxiliary expresses the temporal information. Ultimately it is the auxiliary and the participle together that combine the [+past] and [+perfective] features, but both contribute different pieces of information. Given that the auxiliary encodes primarily a Tense feature in present perfect forms, if

[10]Note that the classification of verbs into different lexical aspect categories cannot always rest on absolute criteria, but must also take into account additional linguistic information, for example whether a verb is used intransitively and can be considered an activity verb (e.g. *Ho mangiato molto, "(I) have eaten a lot"*), or whether the same verb is used transitively with a direct object and can therefore be classed as an accomplishment verb (e.g. *Ho mangiato tutta la torta, "(I) have eaten all the cake"*). The classification of C.'s present perfect verb types has clearly taken into account these variables.

[11] In Italian the use of the simple past (*passato remoto*) in the spoken language has been virtually completed replaced by the present perfect (*passato prossimo*), especially in northern varieties.

Italian children were truly insensitive to Tense features in the beginning, one would expect to find a considerable proportion of auxiliaryless present perfect forms where only aspectual features would be encoded by the use of a past participle. There is some evidence that this is the case for C. The earliest four occurrences of verbal participles with the verbs *cadere, "to fall"*, and *uscire, "to go out"*, are indeed without a tensed auxiliary:

(22) File 7

*CAR: oh caduta.
%eng: oh fallen.

(23) File 7

*CAR: uscita.
%eng: gone out.

As from file 11 (2;3.7), where the next participle form is attested, more than 90% of present perfect forms have a tensed auxiliary. Altogether only 6 forms of the auxiliary *essere, "to be"*, are missing out of a total of 36 obligatory contexts in present perfect forms (16.6%); and 4 tokens of *avere, "to have"*, are missing out of a total of 85 obligatory contexts (4.7%). It seems that C. is aware of the obligatoriness of auxiliary forms in present perfect contexts, and he is also sensitive to the distribution of the two auxiliaries; there are in fact only two substitution errors, in file 20 (3;0.17), where he uses *avere* instead of *essere* with a reflexive verb:

(24) File 20

*CAR: si ha [*] nascosti.
%eng: (s/he) has hidden.
%err: ha = sono $LEX $SUB $AGC

Moreover, 70% of the verbs that appear in the present perfect also appear in at least one other form throughout the period of observation. Although, as already mentioned above, there are no Person contrasts in the present perfect until file 16 (2;10), and despite the fact that a five-way Person contrast in the present perfect is operational only from file 18 (2;11.12), there is some evidence that some Tense contrast can be said to be emerging in the child's grammar around 2;10-2;11.

The analysis of C.'s use of lexical verbs has shown how his acquisition of inflectional morphology proceeds in a piecemeal fashion where Person, Number, Tense, Aspect, and Mood contrasts are added gradually over time, and where, for

example, the acquisition of Person and Number contrasts in the present tense does not immediately carry over to other tenses.

8.6. Conclusions

In this chapter we have presented a detailed analysis of C.'s production of the copula, inflected verb forms, and pronominal object clitics. Although the number of forms is relatively large and varied, and some grammatical contrasts begin to emerge around file 13 (2;5.6) for the copula and inflected lexical verb forms, and around file 14 (2;5.26) for pronominal object clitics, there are also clear indications that much of C.'s learning revolves around a small number of lexical items. The emergence of grammatical contrasts, and hence productive and contrastive realisation of formal features such as Person, Number, and Tense, is a gradual process which is inextricably tied to the expansion of the child's vocabulary to incorporate an ever larger base of forms to reach a critical mass point (see section 9.7.3). By the end of the period of observation (3;0.3), there is some evidence that some contrasts are beginning to become productive, in particular Person contrasts on lexical verbs in the present indicative and in the present perfect, although we do not think that productive and contrastive use of verb forms and clitics has truly reached across-the-board proportions yet. There is therefore reason to be wary in crediting the child with active Agreement and Tense FCs that are required for the checking of formal features.

Chapter 9

A comparative analysis of the emergence of morphosyntax in English and Italian

9.1. Four research questions

Following the analysis of the emergence of C.'s DPs in chapter 6, and the emergence of Agreement and Tense features in English and in Italian in chapters 7 and 8 respectively, we are now in a position to return to the four research questions that were formulated in the introduction:

(1) What is the child's path to the acquisition of formal grammatical features in English and Italian?
(2) To what extent does the emergence of formal grammatical features differ in the two languages?
(3) If there is an observed difference between English and Italian, what factors can be identified to account for it?
(4) Is there evidence for the separate and independent development of the two languages?

The aim of this chapter is to bring together the results presented in chapters 6, 7, and 8 in a unitary and comparative fashion so that a comprehensive picture can be presented of the initial phases of C.'s simultaneous acquisition of two languages.

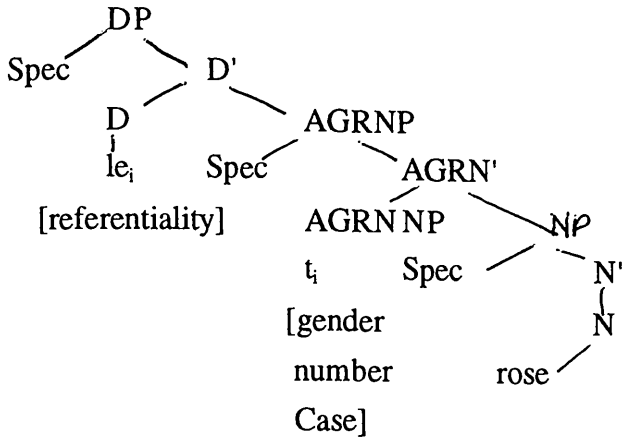
9.2. The nominal system

9.2.1. The syntax of DPs

As outlined in chapter 5, the syntactic representation of nominals involves the complex structure in (5)b:

(5)a Voglio le rose che sono sul tavolo.
"I want the roses that are on the table"

(5)b



The tree diagram in (5)b is a syntactic representation of the object DP *le rose, "the roses"*, in (5)a. The DP in question is clearly referential, and therefore the definite article must be found in D, the higher functional projection associated with the referentiality feature. The DP *le rose* is also marked as being [+plural] and [+feminine], the definite article *le*, and the noun *rose* agree in number and gender by virtue of percolation of the phi features in question from N to AGRN, to D.

The two nominal functional projections, AGRN and D, are both associated with nominal features: D with referentiality, and AGRN with Gender and Number agreement features.[1] An important distinction needs to be made clear between these two sets of nominal features. In Chomsky (1995) phi features on nominals are treated as Interpretable, while phi features on verbs are considered as -Interpretable. In this work we have adopted Tsimpli' and Stavrakaki's (1999) proposal to regard phi features on nominals as -Interpretable. The argument is that agreement features in nominals are realised on all heads of the nominal projection as a result of a N-to-D head Chain formed through percolation, these resumptive features can be erased at LF as -Interpretable (Manzini, 1995). The heads involved in the Chain, by contrast, will survive at LF depending on their remaining feature specification, thus N and D will still be visible at LF, because of their inherent interpretability, but AGRN will not, since it only includes resumptive -Interpretable features that have been erased by the checking process.

[1] In principle AGRN is also associated with Case, the other -Interpretable nominal feature, however since Case is not overtly marked on articles either in English or in Italian, we will ignore this additional feature for the purposes of the present analysis.

This distinction between the Interpretable referentiality feature on the one hand, and -Interpretable agreement features on the other, is crucial for the proposal put forward here that -Interpretable agreement features will be acquired earlier than the Interpretable referentiality feature. The working assumption is that because the interpretation of referentiality lies at the interface between the syntactic and the pragmatic component, children whose initial pragmatic competence is deficient will have difficulties in recognising articles as markers of referentiality. On the other hand the acquisition of agreement features only requires that the child be able to discover the local agreement relation that holds between articles and nouns, a task which, particularly in Italian, is greatly facilitated by the amount of transparent, redundant, consistent and reliable information encoded on noun endings and gender- and number-marked articles.

9.2.2. Predictions for the acquisition of DPs

The syntax of DPs briefly summarised above gives an indication of what is involved in the acquisition of the article system: articles must be identified as markers of referentiality, and phi feature agreement between the article and the noun must become obligatory. In addition, the morphophonological realisation of different articles and nouns must be learnt, and they must be assigned to the appropriate gender and number classes. A bilingual English-Italian child will also have to learn the language-specific distribution of determiners as illustrated in more detailed in section 5.3. In terms of the acquisitional schedule, the predictions made in section 5.10 are repeated here for convenience:

(6)a initially, when children are not yet aware of the relevance of either agreement or referentiality features, they will produce determinerless nominals

(6)b the initial determinerless phase is followed by a stage in which children start to produce their first Article+N combinations, they are almost exclusively singular and lexically restricted to a small number of nouns

(6)c gradually the number of nouns articles appear with increases, and children start to produce Article+N combinations more consistently.

By this time there is still no clear evidence that articles are identified as markers of specificity, but the local head-to-head relation between articles and nouns is appropriately marked by agreement

(6)d the final stage is represented by the identification of the referentiality feature. Not only do children obligatorily satisfy the agreement relations between articles and nouns, they also use definite and indefinite articles appropriately to identify referents appropriately

The predictions in (6) envisage an acquisitional sequence where -Interpretable phi features encoded in a local agreement relationship between the article and the noun emerge earlier than the Interpretable referentiality feature. The claim being made here is that the ability to encode referentiality through the appropriate selection of a definite or indefinite article requires that the child be aware of the specific/non-specific distinction, and that he be sensitive to the given/new distinction, i.e. the child must be in a position to take into account her listener's point of view. Contrary to what initially claimed by Piaget (1962) and Bruner (1966), a number of studies have shown that children are indeed able to make the relevant distinction between specific and non-specific referents from very early on (see Maratsos, 1976; Warden, 1976; Karmiloff-Smith, 1979; Garton, 1983). What children seem to be having more difficulties with is the pragmatic requirement of familiarity. Their failure to take into account the listener's point of view biases children into assuming that what is known to them is also known to the listener, which of course is not necessarily always the case. This error of judgement on the part of the child results in an overproduction of definite articles marked as [+specific], [+familiar], even in those contexts where a [+S; -F] indefinite article would be required.

In sum, the use of articles is cognitively demanding, it requires semantic-pragmatic competence about both the specific/non-specific distinction and the given/new distinction. By contrast, the correct marking of Gender and Number simply requires the child to register the mapping of these features in a local agreement relationship via distributional learning. It is therefore reasonable to predict that -Interpretable agreement features will be acquired relatively early via low-level distributional learning, while the Interpretable specificity feature will require the development of higher level semantic-pragmatic competence.[2]

[2]See López-Ornat (1997) for a similar claim on the acquisition of the morphophonology of articles in child Spanish.

9.2.3. The mapping of features onto functional categories

Assuming a WC approach, the instantiation of the functional projections above NP is subject to the lexical acquisition of the relevant morphophonological properties of articles and nouns. This means that nominal AGR will be instantiated only when the relevant Number and Gender distinctions have become operational in the child's grammar. This translates into obligatory Number and Gender agreement between articles and nouns. In addition, the child must also discover that DPs are associated with the referentiality feature realised by the higher D projection. Unlike the realisation of the AGRN head and its associated phi features which are treated here as - Interpretable features, the instantiation of D requires the acquisition of the referentiality feature, a feature whose interpretation lies at the interface between the syntactic and the pragmatic components. The realisation of referentiality requires the presence of a determiner in the D head and the establishment of an operator D chain: if no D head is available the determiner will remain in AGRN and referentiality will not be marked (see section 5.3).

In sum, what we would expect is that until AGRN and D have emerged, articles will not be produced, there will be an inital determinerless phase where bare nouns are generated in N. A second stage follows where articles start to appear, at this point an AGRN category emerges where articles and nouns agree in Gender, and later Number. For English this stage is only partially visible since definite articles are not overtly marked for Number or Gender, only the indefinite article *a/an* is marked as [+singular], and the indefinite quantifier *some* is marked as [+plural]. For Italian, where the morphophonological marking of Number and Gender follows a very transparent, consistent, and reliable pattern, the prediction is that children will have little difficulty in establishing the relevance of such phi features and the associated AGRN category.[3]

[3]There is empirical evidence that Italian-speaking children have no trouble in the acquisition of Gender and Number distinctions in the article system (cf. Pizzuto & Caselli, 1992; Cipriani et al., 1993). By contrast, in languages such as German where articles also encode case, and there is a high degree of homonymy, there is evidence that the acquisition of the article system is a long and gradual process (Mills, 1985).

9.3. A comparison of the emergence of the article system

9.3.1. A lead-lag pattern in the emergence of DPs

A comparative analysis of C.'s acquisition of definite and indefinite articles reveals a lead-lag pattern between the two languages with Italian several months ahead of English in the production of determiners in obligatory contexts. Figure 1 shows the proportion of determiners in the two languages over time:

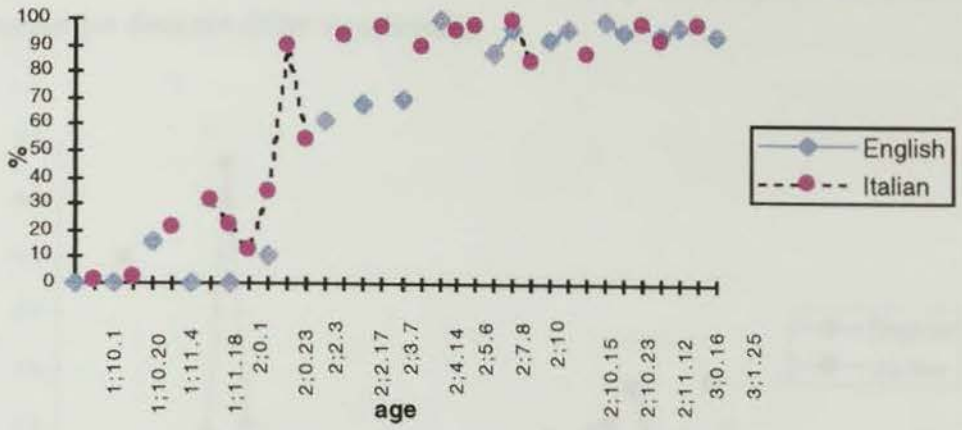


Figure 1. Proportion of articles in obligatory contexts in English and Italian

The first striking difference is that, while in English there are four of the earliest recordings (1;10.1-2;0.1) where there are no determiners at all being produced, in Italian, by contrast, even from the earliest sessions C. produces a number of determiners whose proportion starts to become non-trivial from file 3 (1;11.17) onwards (see also Table I, section 6.2). Therefore, even before determiners reach any significant proportion in obligatory contexts in Italian, there is already some indication that they are starting to be noticed by C.. No such trend is observed in English where for the first month of recording there are no instances of determiners at all, and it is not until file 7 (2;2.12) that determiners start to emerge and are produced with some consistency thereafter.

In Italian C. reaches the 90% target in obligatory contexts in file 8 (2;1.23), while in English it is not until file 10 (2;4.29), some three months later, that he reaches target-like performance. As soon as the child reaches the cut-off acquisition point in English around 2;5, production of determiners in both languages remains fairly stable

around the 90% mark. By this time C. produces both definite and indefinite articles with a large number of nouns in both languages, and there is sufficient overlap between the nouns used with a definite article and the ones used with an indefinite article to assume that the use of the two has become productive.

Another difference between English and Italian nominals is the emergence of Number. As shown in figure 2 below, plural nominals emerge earlier in Italian; as early as file 2 (1;10.27) we find the earliest instances of plural nominals. In English we do not find plural nominals until file 8 (2;2.4), some four months after the first appearance of plural nominals in Italian. It is noticeable, however, that as soon as plural number emerges in English it remains fairly stable, and the proportion of plural nominals in the two languages does not differ significantly.

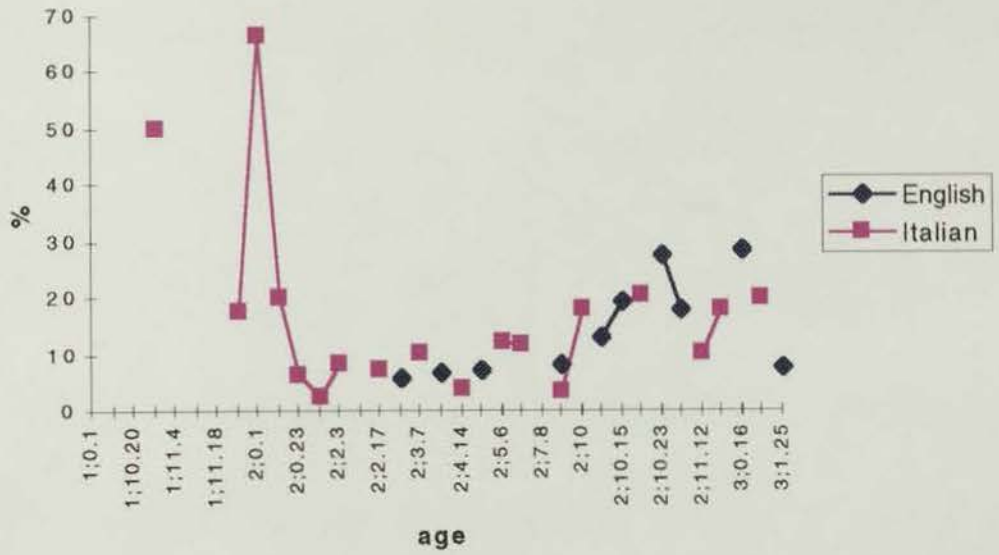


Figure 2. Proportion of plural DPs in English and Italian

By 2;5 we can say that production of determiners in obligatory contexts in English has caught up with production in Italian obligatory contexts, no significant differences between the two languages are noticeable from this point onwards, either in terms of suppliance, or in terms of productivity of use. What is of crucial importance at this point, is to investigate to what extent the production and use of articles differs between the two languages *before* 2;5.

9.3.2. Type/token ratios before 2;5

An analysis of the type/token ratio of nouns in the initial phases of article production shows a revealing difference between the two languages. Figure 2 reports the type/token ratio of nouns appearing with either a definite or an indefinite article before 2;5, when production in obligatory contexts reaches 90% in both languages and remains stable thereafter.

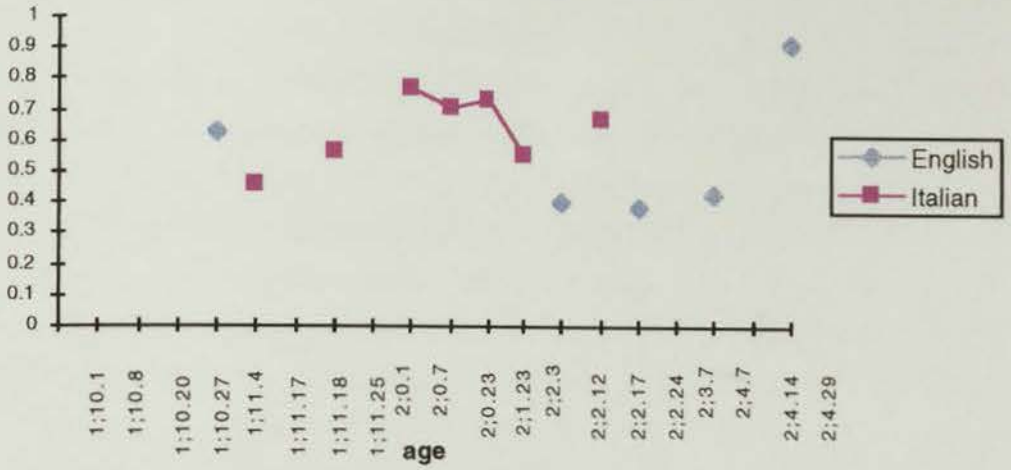


Figure 3. Type/token ratio of nouns in Art+N combinations in English and Italian

Except for a peak of 0.92 in file 10 (2;4.29), the type/token ratio in the earliest English sessions is overall lower than in the earliest Italian sessions. This reflects the limited number of nouns that initially appear with an article in English. In file 7 (2;2.12) for example, out of 13 tokens of Art+N 6 are *a dog*, and 4 *a ball*. In file 8 (2;2.24), 3 Art+N combinations (7 *a dog*, 6 *a cat*, and 5 *a mouse*) account for more than half of all determinate nominals. In the initial stages of article production it is only a handful of nouns that start to appear modified by a determiner in English (*dog, cat, ball, mouse*). There is a clearly observable lexical effect on the emergence of articles in English: C. starts out by relying on a very small set of nouns that can optionally be modified by an article, there is no obvious across the board effect for the use of articles in that they are both lexically restricted, and optional in their use for a protracted period of time. There is an interval of 5 months between the first appearance of articles in English (1;11.4) and the beginning of the stage at which C. uses articles in a consistent and more productive fashion (2;4.29). The realisation of the obligatoriness of articles in English is therefore a relatively long and gradual process. There is no sudden, abrupt change in C.'s production, but rather an ongoing process whereby more and more

nouns are modified by an article. Initially the same nouns that do appear with an article may optionally appear without one, and the article they appear with may be restricted to either a definite or an indefinite. Note that optionality is still residually present even when performance has stabilised well over the 90% cut-off point, however at this point it is limited to a small number of examples, which can be treated as performance errors (see Table VII section 6.4).

If the initial stage in the appearance of articles in English is characterised by a high degree of lexical specificity and optionality, the same does not hold for Italian to the same extent. As noted above, even in the earliest sessions there are a number of determiners (ranging from 1.36% at 1;10.8 to 89.13 at 2;1.23), whereas in English the earliest recordings are characterised by the total absence of determiners (see Table VII, section 6.4). As for lexical restrictions on the nouns that are modified by an article, it is only in file 3 (1;11.17) and 4 (1;11.25), where for the first time we find a non-negligible number of determiner tokens, that we notice a certain bias towards the use of a small number of nouns modified by a determiner. In file 3 (1;11.17) out of 19 Art+N combinations 6 are accounted for by *l'orso, "the bear"*, 4 by *l'altro, "the other"*, 3 by *l'altro poppode, "the other hyppo"*, and 3 by *il gatto, "the cat"*. A similar trend is observed in file 4 (1;11.25) where *il lupo, "the wolf", l'altro orso, "the other bear", il gatto, "the cat"*, and *la scala, "the stairs"* account for 10 of the 14 Art+N combinations. From file 6 (2;0.7) however, the trend is reversed and a larger number of nouns appear modified by an article; it is no longer the case that three or four nouns are found in the vast majority of all of C.'s determinate nominals.

In Italian, like in English, production of articles seems to be kickstarted by the identification of a small number of nouns that will take a definite or an indefinite article. However, unlike in English where this lexically-biased stage lasts for a relatively long period of time (1;11.4-2;4.29), in Italian as early as file 6 (2;0.7) C. starts to extend the use of articles to a large number of familiar and new nouns. In file 6 for example, of the 13 different nouns that are found modified by an article, only 3 are repeated more than once, and in file 7 out of 24 different determinate nominals, only 5 are repeated more than once.

9.3.3. Transparency and closedness of the Italian article system

As the number of noun types in determinate nominals increases, so does the proportion of articles in obligatory contexts. The gradual and constant extension of

article usage to a growing number of noun types is thus accompanied by increasingly better performance in an across-the-board fashion. The transition from the use of articles with a limited number of nouns to across-the-board usage signals the child's growing awareness of articles as obligatory modifiers of nouns. In C.'s case the realisation of the obligatory nature of articles in nominals occurs earlier and faster in Italian than in English. A partial explanation for the earlier age of acquisition and the faster rate of acquisition observed in Italian may be found in the morphophonological nature of Italian articles. Although, like English articles, Italian articles are monosyllabic and unstressed elements, nevertheless, apart from the masculine definite and indefinite articles (*il, "the"*, and *un "a"*), they obey a vowel harmony rule whereby both article and noun end with the same vowel, e.g. *una palla, "a ball", la palla,"the ball", le palle, the balls, i bambini, "the children", uno struzzo,"an ostrich", lo struzzo, "the ostrich", gli struzzi, "the ostriches"*.[4] Despite a few exceptions, the correspondence between the article's vowel and the noun's word-final gender/number-marking vowel is very transparent, predictable, consistent, and reliable.

As suggested by Levy (1983, 1988) for the acquisition of the Hebrew gender system, the transparency, and reliability of the gender/number marking system on Italian articles and nouns must greatly facilitate the acquisition task. The Italian article paradigm, unlike for example the German article paradigm where there is a high degree of syncretism, is remarkably transparent with one-to-one form-function mappings. Transparency and closedness are the key properties that Levy (1996) identifies as responsible for the ease of acquisition of certain formal systems.[5]She argues that opacity of the paradigm will affect the ease with which the paradigm itself will be learnt. In other words, an opaque paradigm with a high number of homonyms and exceptions will be harder to learn than one where there is a transparent and reliable form-function mapping.

In addition to transparency, Levy (1996) also mentions closedness as a crucial property in determining when, and how fast, a paradigm will be acquired. A closed system is such when "its rules and primitives are defined in a way that is internal to the system, with no necessary input from other systems or domains" (Levy, 1996: 83). By this definition the Italian article paradigm as a gender/number marking system is

---

[4]There is also a class of nouns ending in *-e* in the singular and in *-i* in the plural which includes both feminine and masculine nouns, e.g. *il cane/i cani, "the dog/the dogs", la nave/le navi, "the ship/the ships"*. In addition there are a number of exceptions that need to be rote learned: e.g. *la mano/le mani , "the hand/the hands", il pigiama/i pigiama "the pajamas"*, etc.

[5]See section 5.6 for a more indepth review of studies on the acquisition of gender marking.

indeed a closed system, the child will simply need to focus on low-level, transparent morphophonological evidence to figure out how the system works.

In sum, there is evidence for a lead-lag pattern in C.'s two languages, both in the emergence and in the acquisition of articles in obligatory contexts. There is also some evidence that, although an initial degree of lexical specificity in the emergence of articles is present both in English and in Italian, C. does extend article usage to nouns across-the-board earlier in Italian than in English. The greater ease with which articles seem to be acquired in Italian must surely be a reflex of the transparency and the closedness of the Italian article paradigm itself.

9.3.4. The instantiation of nominal functional projections

An important distinction must be noted here in terms of articles as markers of Gender and Number, and articles as markers of referentiality. On the one hand, articles, especially in Italian, signal the Gender/Number specification that is already made visible on the noun. In syntactic terms this translates into the presence of the article in the head of AGRNP, where it checks off the Number and Gender features it shares with the noun in N. As shown in (5)b however, there is a higher functional projection the article must raise to in order to check its referentiality feature, this is relevant only when articles are truly used as marker of referentiality, and children have become sensitive to the given/new distinction that governs the distribution of definite and indefinite articles. The proposal being made here is that when articles start to be used by the child, and Gender and Number are marked correctly, the article is in AGRN; at the same time, if there is no evidence yet that the child is indeed sensitive to the aforementioned referentiality feature and the discourse property of familiarity, then there is no need to assume that the article raises as far as D.

Although the bilingual child will at some point realise that nouns must be modified by an overt determiner where appropriate, and that the distribution of bare nouns is restricted to a limited number of contexts in English and Italian, nevertheless this does not necessarily imply that the child is also aware that definite and indefinite articles are required in complementary distribution to mark referentiality, and encode the listener's familiarity with the referent, or lack thereof. In fact, definite and indefinite articles may initially be used in free variation simply as realisation of the obligatoriness of determiners. The constraints on the distribution of definite vs. indefinite articles may

not be apparent to children as soon as they start to use articles, this kind of pragmatic competence is in all likelihood subject to the emergence of a pragmatic rule, as suggested by Hoekstra & Hyams (1995), and Schaeffer (1997):

(7) Discourse Rule

> Preceding linguistic discourse and knowledge of interlocutor MUST be taken into account.
>
> (Schaeffer, 1997: 72)

If this is the case, children initially have partial representation of DPs where only the functional category that is relevant for the checking of features that are part of the child's grammar needs to be postulated.

In a language like Italian, for example, where there are transparent cues to the Gender/Number marking function of determiners, when articles are present and agree with the noun in N in gender and number, but there is still no clear evidence that the child uses definite and indefinite articles in a pragmatically competent way, the only relevant functional projection that needs to be postulated above NP is AGRNP. The D projection is to be considered part of the child's syntactic representation of nominals only when there is unambiguous distributional evidence that, not only does the child use articles which agree in number and gender with the noun they modify, but that he can use these articles in a way which suggests that he does take into account preceding linguistic discourse and knowledge of his interlocutor. The postulation of a D layer in child grammar is thus subject to the emergence of pragmatic knowledge which interacts with previously established syntactic knowledge about the function of determiners, and articles in particular, as markers of the specific/non-specific distincition and the new/given distinction.

In a similar way in English, there is reason to credit the child with a fully developed DP only when he shows sensitivity both to specificity and familiarity. As in the case of Italian, before such evidence is present, there is no need to assume that the child's nominals have the same syntactic representation as adult nominals. An AGRNP is sufficient to account for the fact that the child can use articles, and that he is sensitive to the number distinction which applies to the distribution of the singular indefinite article *a/an* and to the plural indefinite *some*.

9.4. The transition from predicative to argument DPs

As argued by Longobardi (1994), determiners are operators binding a variable whose range is always the extension of the natural kind denoted by the head noun. In the plural form of common nouns such a range is constituted by members of the extension ; in the singular it is the choice of the determiner which establishes whether the range is constituted by members of the extension of the kind (count interpretation), or by parts of its members (mass interpretation). Longobardi also suggests that the plural form or mass interpretation of common nouns in N is the unmarked realization. In other words, the singular non-mass interpretation of a common noun, as in *a girl*, is simply the result of syntactic agreement with its determiner which designates a single entity. The property of head-to-head syntactic agreement between a determiner in D and a noun in N that "singularises" common nouns should be regarded "as a *marked* lexical peculiarity of certain determiners" (Longobardi, 1994: 634, my emphasis). In other words, singular common nouns, either in English or in Italian, can never receive a definite interpretation in the absence of an overt determiner, their only interpretation can be the unmarked mass interpretation, since bare nouns are not referential, they simply denote a natural kind.

(8)a I ate lion.

(8)b Ho mangiato leone.

The absence of an overt determiner in (8)a and (8)b forces a mass interpretation of the bare noun *lion* and *leone*, with the corresponding meaning of *I ate lion meat.* Similarly in (9)a and (9)b below the bare noun in predicative position can only receive a mass interpretation:

(9)a This is lion.

(9)b Questo è leone.

This property of null determiners to force a mass interpretation is shared in Italian by the partitive article *di + definite determiner*:

(10)a Ho mangiato del leone.
"I have eaten lion"

(10)b Questo è del leone.
"This is lion"

The interpretations of (8)b and (10)a are equivalent in the same way as (9)b and (10)b are.

The correlation between the unmarked mass interpretation and bare nouns has an interesting correlate in child language. As is well known, children have a tendency to omit articles in obligatory contexts in the earliest stages of acquisition, and as a consequence they produce non-target bare nouns where the adult language requires determinate nominals. It is also a known fact that, at least in cultures where object labelling is one of the primary activities parents engage in with their infants, children's earliest utterances tend to be predicative constructions of the type *this is x*, where the *x* slot may be filled by a range of nouns. In the absence of an overt determiner binding the N variable via the formation of a syntactic chain, we propose that children's bare nouns get a sort of direct unmarked mass interpretation. When children first start labelling objects in the *this is x* construction, they are not singling out a particular member of a class as when they use an overt determiner in non-labelling contexts, rather their bare nouns simply denote a natural kind. This suggestion is akin to Maratsos' (1976) proposal to treat indefinite articles in naming contexts as [-S; -F]. The child's principal intention when labelling is to be understood as that of identifying the object as belonging to a class, the object in question can be any member of that class, rather than a particular member.

In a similar fashion, the mass interpretation of bare nouns simply identifies the natural kind denoted by the head noun; there is no intent to single out a specific individual whose property is to be of the natural kind referred to by the noun. When a child says *this is lion*, she is simply signalling what natural kind the noun denotes, she is not singling out a specific lion. In the same way, when indefinite articles start to appear in naming contexts, e.g. *this is a lion*, the child is only referring to any member of a class rather than to an individual lion that belongs to the class. It is obvious that the child's interpretation is not the same as an adult would give, especially because adults know that the mass interpretation of singular nouns is restricted to a relatively small set of nouns. The child's interpretation is clearly non-target, but it is a possible interpretation of DPs in general. Along similar lines, Antelmi (1997) has made the suggestion that monolingual Italian children at the one-word stage use nominals exclusively in the predicative function (*etto bimbo, "this child"*), she argues that the non-referentiality of a predicative nominal makes it a NP and not a DP. This ties in with Longobardi's (1994: 620) claim that "a nominal expression is an argument only if

it is intorduced by a category D". Predicate nominals are not arguments and are therefore not introduced by a category D.

In the adult grammar of English and in Italian bare nouns are also allowed, and Longobardi (1994: 641) argues that they get a default existential interpretation. The assumption is that the D position is filled by a phonetically null determiner, a pure existential operator which binds the noun in the N head, and the result is a semantically unmarked interpretation. Even for bare nouns then, there must be a D position in the DP hosting a determiner, albeit a phonetically null one, which is required to form a syntactic chain with the noun in N so that a semantic interpretation can be given to the DP.

Above we have argued that children's bare nominals are not full adult DPs, therefore no D position is available to host a null D; at the same time we have proposed that bare nouns in child language receive a semantically unmarked existential mass interpretation in the same way as adult's bare nouns do. How can we reconcile the discrepancy between the assumption of the absence of a D projection in child language, and the need for nominals to be assigned a semantic interpretation? The solution proposed here is similar to that originally put forward by Hoekstra & Hyams (1995). In the absence of a D category, or underspecification thereof in Hoekstra & Hyams' account, no syntactic chain can be formed between D and N, hence no grammatical interpretation of nominals is available to the child. Unlike in the adult grammar, in the immature child grammar a direct interpretation is possible whereby a direct link is established between the noun and a referent in the real world without the mediation of a syntactic D operator.

The proposal here is that a qualitative shift will come about in the child grammar when she makes the transition between the "predicative mode", in which nominals are almost exclusively found as predicates in the *this is x* type construction, to the "argument mode", where nominals begin to appear as arguments of verbs, subjects, objects and prepositional objects. The argument mode requires a shift in perspective, the child's focus can no longer be on the naming of the natural kind the noun denotes (bare nouns), or on the identification of an object as any member of a class (naming), she now needs to identify a particular individual as the subject or the object of the verb that she uses to describe a specific event, state of mind, or activity. In other words, the child must realise the indexical function of determiners. When the child says something like *the lion runs*, he is obviously picking out a specific individual lion of

which he predicates that it runs, and he must signal the fact that he is talking about a specific lion by using a determiner. The child must learn about [+referential] DPs and [-referential] Dps. The implication however is not necessarily that the child will also choose an appropriate definite vs. indefinite determiner depending on the listener's familiarity with the referent in question. This is an additional piece of pragmatic knowledge that may take some time to develop as reported on the vast literature on egocentric errors (see section 5.5)

In C.'s data we find precisely such a qualitatively shift between predicative and argument mode in the use of bare nouns and [-S;-F] indefinite articles in predicate nominals on the one hand, and the use of definite articles [+S;+F] and [+S;-F] indefinite articles in argument nominals on the other. In both languages there is a fairly abrupt transition from an initial phase characterised by an overwhelming proportion of predicate nominals, to a second phase in which the large majority of nominals are found in argument position. In both languages the shift from a higher proportion of predicate nominals to a majority of argument nominals is accompanied by an increase in the production of articles in obligatory contexts, and in the emergence of verbs. Figure 4 below reports the proportion of predicate and argument nominals for English and Italian:

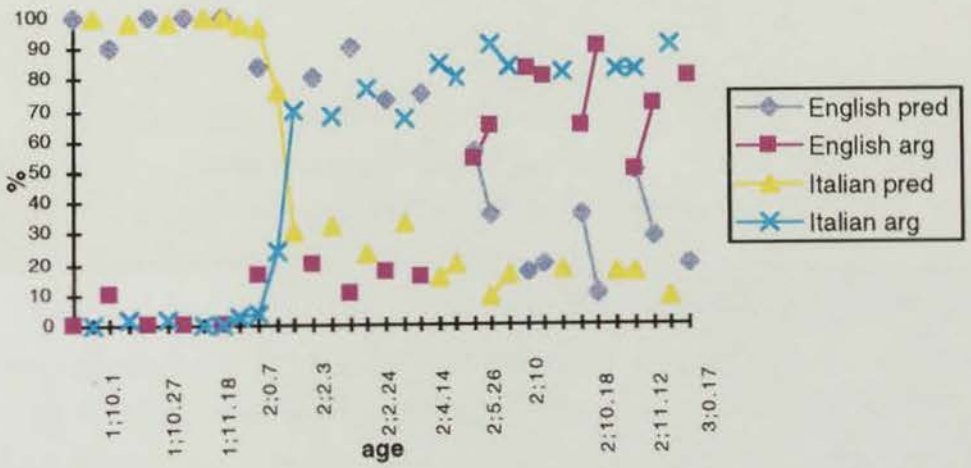


Figure 4. Proportion of English and Italian nominals in predicate and argument position

As shown in Figure 4, in Italian until file 9 (2;2.3) the majority of nominals are found in predicate position. From file 9 onwards the trend is reversed, and 70% and over of nominals are found in argument position. Just before this qualitative shift in the

distribution of nominals, we also observe that in file 8 (2;1.23) for the first time the production of determiners in obligatory contexts reaches 89.13%. The realisation of the obligatoriness of determiners is clearly paralleled by the increase of nominals in argument positions. This is indicative of a correspondence between the two phenomena, although we still have to verify whether there is indeed a correlation between the two, i.e. whether it is the case that determiners appear with argument nominals, and bare nouns are found in predicate nominals. When one looks at the proportion of residual determinerless nominals after 2;2.3, it is in fact the case that the vast majority are actually found in predicate position (96%) and not in argument position (4%).

The majority of C.'s Italian determinate nominals in argument position are singular and specific the largest group of determinate nominals is in fact accounted for by nominals with a definite article [+S;+F], and nominals with an indefinite article [+S;-F] in non-naming contexts (see Figure 2, section 6.2). By the time verbs start to emerge and singular, specific arguments are expressed overtly, the child is becoming aware of the need to make the referentiality feature on the noun visible, hence the emergence of definite and indefinite articles. As already mentioned above, although referentiality may be marked visibly by the use of an overt article by this stage, the child may still have incomplete mastery of the relevance of the familiarity [F] feature, and may still commit a number of egocentric errors. C. does in fact make a number of egocentric errors with Italian definite articles, although the average error rate is not high, around 3.37%.

A similar trend is observed for the English data. In file 11 (2;7.8) the proportion of argument nominals exceeds the proportion of predicate nominals for the first time. The gap between the two nominal positions remains considerable and constant, with the exception of file 17 (3;0.3) where there is a 50% proportion of each. As in Italian, the shift in the predominace of predicate nominals follows the stabilisation of determiner production in obligatory contexts around the 90% mark. In file 10 (2;4.29), immediately before the transition point between predicate and argument nominals, we find a 100% suppliance of determiners in obligatory contexts. Once again the correlation between argument nominals and obligatory determiners is fairly clear. After 2;7.8, only 7% of determinerless nominals are found in argument position.

Despite a very similar pattern in the emergence of obligatory determiners in the transition between the initial predominance of predicate nominals and the subsequent majority of argument nominals, there is a non-negligible difference between English and Italian with respect to the time of occurrence of such transition. In Italian as early as file 9 (2;2.3) C. shifts from the predicate mode to the argument mode, while the same transition does not take place in English until file 11 (2;7.8), some five months later. The main variable responsible for such a lead-lag pattern between the two languages is the emergence of verbs other than the copula in the child's lexicon. It is only when C. starts to use verbs that the need to express arguments becomes manifest; before that his preferred mode of interaction is naming using constructions of the type *this is x*, or formulaic utterances such as *what's that?* or *where is it?*. In Italian verbs emerge earlier than in English and this has obvious consequences for the expression of verbal arguments (see chapter 7 and chapter 8).

9.5. A crosslinguistic parallel in the emergence of DPs

In sum, as far as the emergence of DPs in C.'s two languages is concerned, we have argued for an inital NP stage where determiners are completely absent. This phase lasts approximately until 1;11.25 in Italian, when the first non-negligible proportion of determiners is observed for the first time, and until 2;2.12 in English. In this first phase there is no reason to postulate a higher functional projection above NP, bare nouns are the norm in C.'s nominal production, there is no indication that the child has discovered the syntactic function of determiners and their pragmatic distribution. This initial determinerless stage is also characterised by what we have called here the "predicate mode", the overwhelming majority of C.'s nominals at this stage are found in predicative position in constructions of the type *this (is) x, questo (è) x,* where *x* is a slot that can be filled by a range of (mostly singular) nouns.

The beginning of a second stage coincides with the emerge of a significant proportion of determiners in obligatory contexts (around 90% and over), and the shift from the "predicate mode" to the "argument mode". In this second phase, which begins around 2;2 in Italian and around 2;7 in English, verbs start to emerge, hence the need to express verbal arguments overtly. Naming is no longer C.'s principal mode of linguistic interaction, he now turns to the world around him and starts to describe states, and events that involve specific, individual referents.

When definite and indefinite articles start to appear in Italian they are overtly marked for gender and number, and they agree with the head noun they modify as

required in the adult grammar. The presence of articles that share the same gender/number features of the head noun in N justifies the postulation of one functional category above NP: AGRNP, whose head AGRN is occupied by the determiner. Following the proposal by Tsimpli and Stavrakaki (1999), we have argued here for the -Interpretability of resumptive phi features in AGRN, and for the interpretability of the specificity feature checked off by a higher functional D projection above AGRNP.

The postulation of an AGRNP is also required for the English data when there is evidence that determiners are supplied in at least 90% of obligatory contexts, and when C. shows awareness of the relevance of the Number feature on the indefinite article *a/an* which is specified as [+singular].

If there is reason to credit C. with a higher functional projection above NP when determiners are supplied in 90% or more of obligatory contexts, and gender and number features agree appropriately with the noun in N, we believe that a DP projection above AGRNP is not warranted by the data. There is simply not sufficient empirical evidence to be sure that C. is indeed aware of the pragmatic given/new distinction which governs the distribution of definite and indefinite articles. Although the proportion of egocentric errors is low (an average of 3.37% for Italian, and 4.78% for English), the interactional context of the recording sessions is generally heavily biased towards the deictic use of the definite article, rather than the anaphoric use which requires that the child take into account previous linguistic discourse and knowledge of the interlocutor. There are only a few occasions in which C. is in a position where he can suspect that his interlocutor is indeed not familiar with the referent he has introduced with a definite article. By and large, either he refers to entities that are physically present and can be seen by all participants in the conversation, or he refers to well-known characters in books and fairy tales, archetypal characters which are [+S;+F] by default (Bennet-Kastor, 1983). Additional experimental evidence would be required to settle the matter with a greater degree of confidence; the naturalistic data in this corpus, as is often the case for spontaneous corpora, does not provide the ideal conditions for testing specific hypotheses such as the mastery of the given/new distinction.

Although C. seems to be following a similar developmental path in the acquisition of DPs in Italian and in English, there is an obvious lead-lag patter with Italian leading by an average of five months, both in the suppliance of determiners in

90% or more of obligatory contexts, and in the shift from the predicative to the argument mode. In addition we also observed a qualitative difference in the degree in which lexical specificity determines the production of determiners, especially in the initial phase. If it is true that both in Italian and in English C. relies on a small set of nouns that appear with determiners, this lexical bias seems to last longer for English than for Italian where he extends the use of determiners to nouns across the board at a much earlier age (2;0.7 in Italian, vs. 2;4.29 in English).

In conclusion, we have argued here for the independent acquisition of an AGRN category above NP in C.'s English and Italian nominals. Because of the lack of sufficient empirical evidence, we have tentatively proposed that no D category is present by the end of the data collection period in either English or Italian. Although there is evidence that C. can discriminate between [+specific] and [-specific] referents, nevertheless there is no clear proof that he can indeed take into account previous linguistic discourse and knowledge of the interlocutor as required by a pragmatically competent speaker.

In the following sections we will contrast and compare C.'s performance in the verbal domain to assess if and when he can be credited with knowledge of Agreement and Tense features resulting in the instantiation of AGR and T.

9.6. The emergence of AGR and T

9.6.1. Morphosyntactic diagnostics

From a WC perspective, a child's grammar can be said to include the AGR and T functional projections when there is morphosyntactic evidence for the productive and contrastive use of verbs. Both morphological and distributional information is needed to ascertain to what extent the child's use of verbs truly reflects the instantiation of a functional category. The child must show that he has command not only of some of the inflections of the verbal paradigm, but that he has knowledge of the abstract notion of subject-verb agreement across the board in the present and in at least one other tense. In order to assume that Person and Number are formal grammatical features that have a syntactic representation in the child's grammar, one must find evidence that person and number inflections are used productively and contrastively by the child. The measure of productive contrast for an inflection at a certain point in time is determined by the occurrence of at least two verb types for a given inflection, with at

least one of the two (or more) verb types with that inflection which are also inflected for another person, number, tense or mood. For example, evidence that a child can use the 2 p.s. present indicative inflection productively in Italian would be given by the occurrence, at a given time point, of at least two verb types inflected for 2 p.s. present indicative, e.g. *mangi, "(you) eat"* and *dormi "(you) sleep"*. In addition, an indication of contrastive use would be supported by the cooccurrence of at least one of the two verbs inflected for 2 p.s. present indicative with another person inflection, e.g. 1 p.s. present indicative inflection *dormo, "(I) sleep"*.

Morphological evidence that the child has grammaticalised the notion of Tense is found in the ability to distinguish speech time from event time. Only when the child can talk about an event, a state, or an activity that is not simultaneous to the speech time of her utterance, can we infer that a grammatical notion of Tense is beginning to emerge. In other words, the child must show productive and contrastive use of Tense by contrasting the present, where speech time and event time coincide, with the past and/or the future where speech time and event time are distinct.

Syntactic distributional evidence for the existence of AGR and T is to be found in the child's use of word order. In English, for example, the presence of auxiliaries and modals to the left of negation is an indication that some functional position outside the VP must be present. Subject-auxiliary inversion in English matrix questions is also typically used as evidence for a higher Complementiser projection. In Italian, the use of a syntactic diagnostics for verb movement past negation is somewhat more difficult to apply since there is no obligatory postverbal negator like French *pas* that could be used as a marker of verb movement across NEGP to a higher functional projection.

In the following sections we will apply the aforementioned criteria of productive and contrastive use of verbal morphology, and where possible we will use syntactic diagnostics to assess to what extent C. can be said to have AGR and T in his two languages. As we did for the analysis of nominals in section 9.3, we will review the findings in the two languages in a comparative way to assess to what extent the acquisition proceeds in a parallel fashion in English and Italian, and to verify whether there are any significant crosslinguistic differences.

9.6.2. The copula as an Agreement and Tense marker

Jaeggli & Hyams (1987) suggest that the English copula *be* is an expletive verb inserted into the derivation to carry tense and agreement features, and we add here that an analogous proposal can be made for the Italian copula *essere*. The copula in predicative constructions does not express any particular semantic content of its own, it merely spells out the finiteness of the clause by making Agreement and Tense features morphologically visible. Because of the reduced semantic content of the copula, language acquisition researchers working in a number of different languages have noted time and again that children often omit it in the initial phases of their linguistic development. Besides the lack of semantic salience, another fact predicts the initial omission of the copula: its absence is to be expected on the assumption adopted in this work that the Agreement and Tense features with which the copula is associated, are not actively represented in child language from the onset, but become part of the child's grammar in a gradual, lexically-driven fashion. If AGR and T are not part of the child's grammar from the onset, it is expected that the copula will be omitted. Moreover, it has also been noted that when the copula first appears, it is typically restricted to a limited number of 3 p.s. contexts in semiformulaic utterances of the type *that's x, it's x* (see Radford, 1990: 164-168). These earliest uses of the copula are not to be taken as positive evidence that the child is sensitive to the relevance of Agreement and Tense features, at least not until there is an indication that usage of the forms is productive and contrastive, and that Agreement and Tense features are also morphologically spelled out on lexical verbs.

The following section will compare the findings on C.'s use of the copula in English and Italian presented in section 7.2 and 8.2 respectively.

9.6.3. The use of the copula in English and Italian

When comparing C.'s performance on the copula in the two languages, the first striking difference is in the level of suppliance in obligatory contexts. While in English there are virtually no tokens of present tense copula until file 9 (2;4.7), with the exception of 2 tokens in file 3 (1;11.4), in Italian, tokens of the present tense copula appear as early as file 2 (1;10.27). Moreover, in Italian the percentage of copula tokens in obligatory contexts stabilises in file 6 (2;0.7) and file 7 (2;0.23) at around 67% and from file 8 (2;1.23) it reaches a fairly stable 90% proportion in obligatory contexts. In

from file 10 (2;4.29) onwards, and it never reaches a stable proportion of 90% and over. Figure 5 below charts the development of copula forms in obligatory contexts (singular and plural tokens have been collapsed).

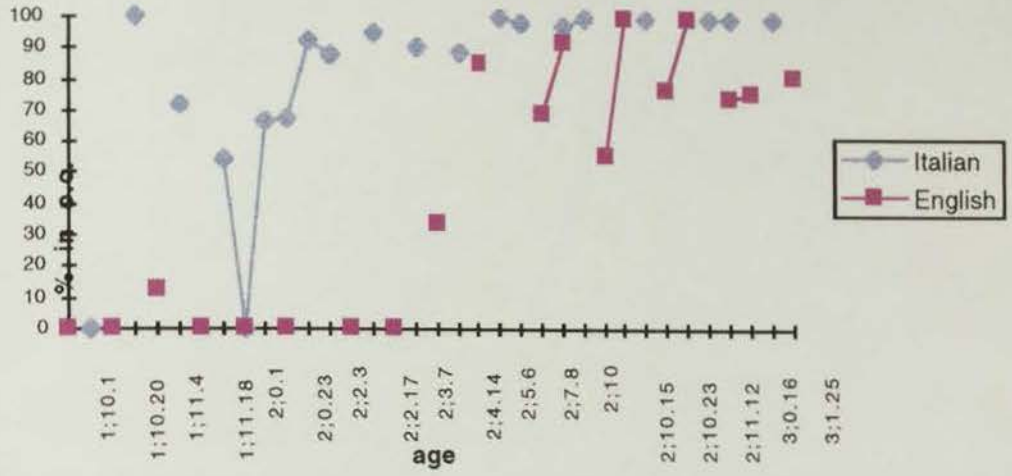


Figure 5. Proportion of present tense copula forms in obligatory contexts in English and Italian

Notwithstanding the quantitative difference observed in the proportion of copula forms in obligatory contexts in English and Italian, there are a number of qualitative similarities in the developmental pattern of this expletive verb. Firstly, in both languages the only two contexts that have any significant representation are 3 p.s. and 3 p.p., with plural contexts emerging later than singular ones (see Table I, section 7.2 for the English data, and Table III in section 8.2.1 for the Italian data). In English the first 3 p.p. copula appears in file 12 (2;9.6), while in Italian we have the first occurrence of a 3 p.p. copula in file 11 (2;3.7), once again a lead-lag pattern with Italian approximately 6 months ahead.

As we saw in detail in section 7.2, the overrepresentation of 3 p. contexts for the English copula was accounted for by C.'s inclination to name objects in the *that's (a) x,* or *it's (a) x* constructions. To a certain extent this is also the case in Italian, although in Italian C. also has the null subject option in naming contexts, where he can simply use the copula followed by a noun or an adjective, without a dummy subject such as English *that* or *it:*

(11) File 4

*CAR: è un lupo.
%eng: (it)is a wolf.

Unlike in English, where the copula is exclusively found in predicative naming constructions, in Italian the copula also figures in existential constructions and in conjunction with a locative:

(12) File9

*CAR: c'è la volpe.
%eng: there is the fox.

(13) *CAR: è lì.
%eng: (it) is there.

In sum, although in both languages the almost exclusive appearance of the copula in 3 p. contexts is related to the naming function it is involved in in predicative constructions, in Italian there is some evidence of a somewhat more flexible use in non-naming contexts such as those exemplified by (12) and (13).

Because of the limited knowledge of the copula paradigm, and because of the extremely narrow range of subjects the copula takes in C.'s English, it was concluded in section 7.2 that the copula is nothing more than an element in a productive positional pattern, rather than the morphological spell out of Agreement and Tense features. Similar conclusions were also drawn for the use of the Italian copula, in Italian too copula forms are limited to 3 p. contexts; there is some sort of Number contrast emerging, but it is limited to 3 p., there is no evidence of an across-the-board Number contrast between persons other than 3 p.. Differently from English where the copula is only ever found in present tense contexts, in Italian from file 13 (2;5.6) onwards there is some contrastive use of the copula in imperfect tense contexts (see Table VII in section 7.2). Here too, all but 2 forms are 3 p., and specifically 3 p.s., therefore the emerging Tense contrast is largely restricted to 3 p.s..

Overall the evidence reviewed so far for the copula in English and in Italian is not robust enough to conclude that, at this stage, the emergent use of this expletive verb in C.'s grammar represents something other than the realisation of a slot in a semiproductive positional pattern. It is also interesting to note that although the supposed reduced semantic content of the copula is virtually the same in the two

languages, C.'s performance in obligatory contexts is worse in English than it is in Italian. It may well be that additional phonological and articulatory reasons also play a role in the production of the English copula. Unlike the 3 p.s. Italian copula *è* [e], the English 3 p.s. copula *is* [ ], when suffixed to a pronoun like *that* and *it* results in a cluster including a dental and a sibilant [ts], a sound which poses articulatory problems to young children. As we will see in the following section where we compare the emergence of Agreement and Tense features in lexical verbs, there is reason to believe that the conclusions drawn for the morphosyntactic role of the copula are in keeping with more general considerations on the emergence of verbal morphology on lexical verbs and on the use of auxiliaries and modals.

9.7. The marking of grammatical contrasts on lexical verbs

9.7.1. Methodological considerations

As is well-known, English and Italian vary substantially in the extent to which subject-verb agreement is morphologically marked on lexical verbs. In the present tense, English only marks a Number distinction, *-s* vs. *ø*, while Italian, a Person-marking language, marks each of the six different person/number combinations with a unique ending (see Table VIII, section 8.3). Person/number contrasts are also uniquely marked in Italian in all other non-periphrastic tenses and moods, e.g. imperfect, preterite, present and past subjunctive, present and past conditional, and simple future, whereas in the English preterit, no Number distinctions are marked at all and only Past is marked in the form of an *-ed* suffix for regular verbs, or a suppletive form for irregulars. The typological difference between these two languages with respect to the morphological richness of the agreement paradigm has obvious implications for language acquisition. On the one hand, Italian-speaking children have to master a larger array of morphological contrasts than their English-speaking peers, on the other, their task is greatly facilitated by the transparency of the morphological paradigm.

Numerous studies on the acquisition of English morphology starting from the seminal work by Cazden (1968), Brown (1973), and de Villiers & de Villiers (1976) to a host of more recent studies (Wexler, 1994; Hyams, 1992; 1996, among others) have documented the relatively protracted period of optionality during which children alternate between infinitival forms, which in English are indistinguishable from the bare stem, and inflected *-s* forms. By contrast, studies on the acquisition of Italian morphology have reported that from the onset children produce a variety of inflected

forms, a high proportion of which also correctly agree with their subject in person and number (Hyams, 1986, 1992; Pizzuto & Caselli, 1992; Guasti 1993/94; Cipriani et al., 1993; Antelmi, 1997).

Some of these studies in the acquisition of Italian morphology have taken the early presence of inflected forms in children's speech as evidence that subject-verb agreeement is mastered earlier in a richly inflected language like Italian than in a poorly inflected language like English. This line of argumentation has often been used by detractors of maturational explanations for the emergence of FCs to dismiss the hypothesis that FCs are subject to a biological maturational schedule. If this were indeed the case, the argument goes, then we should not expect crosslinguistic differences in the emergence of FCs. Admittedly there may be different "chronological stages" due to individual differences, as have been observed in the timing of acquisition within a single language, whereby children of the same chronological age may different greatly with respect to the grammatical stage they are at, but there should not be different "grammatical stages". In other words, if there is a maturational schedule which is at least partly responsible for the acquisition of FCs, we would expect to find in all languages of the world a grammatical no-FCs stage like the one proposed for English by Guilfoyle & Noonan (1988, 1992), Lebeaux (1988) and Radford (1990, 1992, 1994, 1996), and for Swedish by Platzack (1992).

Contrary to the predictions made by the maturational hypothesis, a number of researchers working in languages with richer inflectional system than English (e.g. German, French, Italian, Dutch), have found no indication that such a grammatical no-FCs exists. Their common argument is that in these languages there is morphological and syntactic evidence for the existence of FCs outside VP from the earliest stages of acquisition, therefore the no-FCs stage cannot be maintained as a developmental universal.

The kind of morphological evidence that is normally adduced for the acquisition of subject-verb agreement, and hence for the presence of an AGR category, is 90% suppliance in obligatory contexts. Clahsen, Eisenbeiss & Penke (1996: 143) state that "[they] assume that the paradigm of subject-verb agreement is acquired if more than 90% of the forms are used correctly in obligatory contexts". Without any further qualifications this criterion based on low or residual error rates is clearly not adequate to identify productive and contrastive use of inflected forms. We have argued throughout this work that only clear evidence that the child uses verb forms not only

productively, but also *contrastively* should be used to argue for the existence of an abstract across-the-board notion of subject-verb agreement, and hence for the instantiation of an AGR category. Without any clear indication of this kind of productive and contrastive use of verb forms, no definitive conclusions can be drawn with respect to the grammatical representation of these inflected forms.

It is also unfortunate that much of the current research on the acquisition of FCs relies so heavily on the significance of error rates, taking a low error rate as an indication of mastery. For example, the fact that Italian-speaking children produce inflected verb forms as as soon as they start producing their first verbs, and never produce bare stems is hardly surprising given that by the time they start speaking they must surely be sensitive to the phonotactic constraints of their language which disallow such bare stems.[6] In itself then, the fact that Italian, Spanish, or Catalan speaking children always produce inflected verb forms is no more than a reflex of children's sensitivity to the phonotactics of their language, and in principle it should not be taken as direct evidence of a more abstract kind of knowledge.[7]

If one goes beyond the *prima facie* evidence that children acquiring languages with rich Agreement systems produce a variety of inflected forms, and if one applies the standard productivity and contrastivity criteria originally proposed by Cazden (1968) and Brown (1973) for English, it turns out that these children are not that different from their peers acquiring a less richly inflected language such as English or Swedish. Although at first sight children acquiring Italian or Spanish appear to be using a relatively large number of inflected forms, thus showing a seemingly sophisticated knowledge of the agreement paradigm, at a closer inspection one finds that, although children use a variety of verbs in each of a number of forms, there is little evidence to credit them with a contrastive command of anything. In a very detailed study of the acquisition of Spanish morphology in two children between 1;6 and 2;6, Gathercole, Sebastián & Soto (1999: 29) conclude that

> there is little evidence here of any broad-based knowledge on the part of Maher and María. Instead, several aspects of the data point to piecemeal acquisition - one form per verb at initial stages; the very gradual addition of verbs to elements of the paradigm even when those elements become contrastive; the lack of generality of application of person across tenses even when they enter one tense and become contrastive in it.

---

[6]See Jusczyk, Cutler & Redanz (1993), and Jusczyk, Friederici, Wessels & Svenkund (1993) for evidence of infants' early sensitivity to phonotactic constraints of their language.

[7]Hyams' (1986) postulation of a stem parameter to account for the lack of uninflected forms in Italian child grammar is therefore unmotivated.

9.7.2. Lexically-specific patterns in C.'s use of inflected forms

As was shown in chapter 7 and 8 for C.'s acquisition of verbal morphosyntax, there is evidence for a similar lexically-specific pattern in the emergence of verbs in English and Italian. In English there is an extremely limited lexically-specific use of Number and Tense inflectional morphology, and the use of modals is heavily constrained by the verbal complements they may take (see sections 7.4 and 7.6). Even in Italian, where the number of inflected forms is much larger and more precocious than in English, a fine-grained verb-by-verb analysis revealed a trend towards verb-specific learning rather than paradigm-based learning. Contrastive use of Person, Number, Tense, Aspect and Mood clusters around a small number of verbs for which an increasing number of contrasts becomes available, and there is no clear indication that by the end of the data collection period C. is able to extend contrastive use acquired for one verb type to other verb types in an across-the-board fashion.

The criteria of productivity used in the analysis allowed us to identify cut-off points at which we concluded that a contrast had become active in C.'s grammar. Table XVII, section 8.3.5, illustrates the emergence of a number of formal contrastive uses in terms of Number, Person, Tense, Aspect and Mood. The picture that emerges is one where a number of paradigmatic contrasts gradually appear over time. There is however an important aspect of C.'s acquisition of verbal contrasts that Table XVII does not do justice to: the emergence of these contrasts is still tied to the acquisition of new verb forms for a selected and small number of verb types. The longitudinal verb-by-verb analysis in Appendix 3, presents a more faithful picture of the emergence of contrastive use where we can see how the acquisition of contrastive verb forms is centred around a handful of verb types (see also Table XV, section 8.3.4.). For example the verb *fare, "to do/to make"* is the verb with the single largest number of contrastive forms (N= 13), and it is also the only verb for which C. has a productive and contrastive use of all the person/number combinations in the present tense.

The overall picture that emerges from this overview of various approaches to the acquisition of morphology, and its implications for the realisation of FCs in ECG, is one where crosslinguistic differences in the emergence of inflected forms should be evaluated with greater attention to productivity and contrastivity criteria. The conclusion that children learning more highly inflected languages, like Italian, acquire the morphosyntax of their language earlier than their counterparts learning a language like English, is at best a hasty one. We have shown that if one applies standard productivity

criteria to children's early production of inflected forms, one finds that there are indeed similar grammatical stages across languages. If despite the appearance of a variety of inflected verb forms one can demonstrate that they are not used productively and contrastively, the claim that their use reflects the instantiation of an AGR, and possibly a T category, in the child's grammar is considerably weakened. By this rationale, in the absence of concrete evidence that the notion of subject-verb agreement is operational across the board, it is reasonable to argue that even the ECG of children learning languages with a rich and transparent agreement system go through an initial stage where AGR and T are missing. Similar claims have been proposed by Tsimpli for Greek and a number of highly inflected languages (Tsimpli, 1991; 1996), and Antelmi for Italian (1997).

While the no-FCs stage is more obviously visible in a poorly inflected language like English where children are known to go through a relatively long initial period in which morphosyntactic Agreement and Tense markers are either totally absent or optional, children learning a language like Italian are deceptively good at morphology as soon as they start producing one-word utterances. However, the fact that they do produce inflected forms from very early on is an inevitable consequence of the fact that verb stems cannot stand alone as well-formed morphophonological units of representation in the language. In other words, the seemingly sophisticated and advanced knowledge of the agreement paradigm in Italian children's ECG that has led some researchers to postulate the existence an AGR category from the onset of acquisition, must be subjected to stricter productivity criteria to ascertain to what extent an abstract grammatical representation of Agreement is actually present. Of course, the possibility still remains that, given the greater transparency of the Italian agreement paradigm, children learning this language may actually take less time in figuring out the subject-verb agreement paradigm and hence the no-AGR stage may be shorter than for English-speaking children. In the next section such a possibility is entertained in terms of the "critical mass hypothesis" where a correlation is proposed between verb lexicon size and acquisition of productive use of morphology.

9.7.3. The "critical mass hypothesis"

Given the significant lexically-driven aspect of C.'s acquisition of verb morphosyntax, and the general inability to extend newly acquired contrasts to a large number of verbs across the paradigm, we have concluded that the instantiation of FCs

which are required to check the formal features being acquired by the child is obviously an emergent process which is characterised by the gradual acquisition of an ever larger verbal lexicon. It is only when the number of verb types and verb forms in the child's vocabulary reaches a "critical mass" that a category can be said to be fully active and part of the child's grammar. The "critical mass hypothesis" as proposed by Marchman & Bates (1994) suggests exactly such a correlation between lexical learning and grammatical acquisition. Using information from the administration of the MacArthur Communicative Development Inventory, a parental report questionnaire, to 1130 American English-speaking children between 1;4 and 2;6, Marchman & Bates found a powerful non-linear correlation between the increase of children's verb lexicon and the onset of past tense overregularisation errors. Arguing against a dual-mechanism for the acquisition of past tense inflectional morphology (Marcus, Ullman, Pinker, Hollander, Rosen & Xu (1992), they advocate a single learning mechanism where the emergence of overregularisation errors does not reflect the appearance of a new, qualitatively different rule-based mechanism for past tense production. The hypothesis that Marchman & Bates defend is one where lexical and morphological acquisition are inextricably tied, and where the achievement of a lexical "critical mass" provides the child with a dataset which is large enough and representative enough to extract general (morphological) patterns.[8]

This lexicalist approach is in line with traditionally lexicon-oriented syntactic models such as Lexical Functional Grammar (Bresnan, 1982, in press), Head-Driven Phrase Structure Grammar (Pollard & Sag, 1994), and Construction Grammar (Fillmore, Kay & O'Connor, 1988; Goldberg, 1995), and with recent developments in Chomsky's (1995) Minimalist Program (MP). In the MP more and more of the explanatory work that was previously handled by the grammar has been moved into the lexicon. The richness and diversity of linguistic forms within any particular language are now captured almost entirely by the lexicon which includes "the information required for further computation - in particular, for the operations of the phonological component (including morphology, we assume)" (Chomsky, 1995: 239). If grammar is indeed an inherent part of the lexicon, as is the current view in many syntactic theories, we would expect just such a powerful relationship between grammatical and lexical development.

[8]See also Bates & Goodman (1997, 1999) for an extension of this proposal to atypical populations, and Jones & Conti-Ramsden (1997) for an application of the critical mass hypothesis to the explanation of the delayed acquisition of verbal morphosyntax in SLI children.

If the "critical mass hypothesis" makes a meaningful contribution to the explanation of how children acquire productive use of verbal morphology, we propose here that its logical consequence is that it is possible to think of crosslinguistic differences in the rate and age of acquisition of verbal inflectional morphology in terms of different thresholds at which the critical mass point is reached for different languages (and possibly different individuals). Marchman & Bates have observed that, irrespective of chronological age, children reached the critical mass point for the onset of overgeneralisation errors in English past tense morphology, when the verbs in their vocabulary reached the cut-off point of 70 types. In a comparative study of verb use by children with Specific Language Impairment (SLI) and their younger siblings, Jones and Conti-Ramsden (1997) report that although the size of the verb lexicons of the SLI children and of their younger siblings was not dissimilar, the SLI children did not use the verbs in their lexicon as widely and as flexibly as their normal controls. Although the children with SLI had reached a similar critical mass in their verb learning compared with their normal developing siblings, nevertheless their use of verbal morphology was not as extensive as their brothers' and sisters'. Jones & Conti-Ramsden (1997: 180, my emphasis) therefore propose that "children with SLI require a *larger vocabulary* than normal language learning children in order to reach an 'SLI critical mass'".

In a similar vein, the suggestion being made here is that children learning typologically different languages will have different verb vocabulary sizes when they reach the critical mass point. Children learning languages where morphological information is encoded in rich and transparent paradigm systems are likely to need a smaller number of verb types before their reach the critical mass threshold. By contrast, children learning a language like English, where the verbal inflectional paradigm is not so transparent and consistent, will need to acquire a larger number of verbs before they can reach the critical mass point at which they can start making morphosyntactic generalisations.

In this view, the supposed advantage of Italian-speaking children over English-speaking children in the acquisition of verbal morphology is accounted for by a lower threshold after which the former can start abstracting general morphological patterns of subject-verb agreement. If the Italian inflectional system is more transparent than the English system, it is reasonable to assume that children learning Italian will need to learn proportionally fewer verb types before the pattern begins to emerge. At this point this is only a tentative speculation, the data from C.'s corpus do not provide direct

evidence for this differential critical mass hypothesis, since by the end of the data collection period English inflectional morphology is still a long way away from being mastered, no overgeneralisation errors are attested, and in Italian mastery of the paradigm is still limited to a relatively small number of verbs. Ideally one should test the validity of this proposal by comparing verb vocabulary sizes of children that can be said to have acquired productive and contrastive use of verb inflectional morphology in different languages. If this hypothesis is along the right lines, we would expect to find smaller verb vocabulary sizes in Italian children that have productive and contrastive knowledge of the subject-verb paradigm, when compared to English-speaking children who can be credited with similar verbal morphosyntactic knowledge.

In the following section the issues of productivity, contrastivity, and the relationship between verb vocabulary size and control of the subject-verb agreement paradigm are examined in more detail in C.'s English and Italian data. The conclusion drawn at the end of the analysis points towards a separate but parallel morphosyntactic development in the two languages.

9.8. Evidence of productivity of AGR and T in C.'s English and Italian

9.8.1. The status of AGR and T in English

The status of AGR and T in C.'s English is characterised by a general absence of Agreement and Tense markers in obligatory contexts. The only two AGR/T markers that have any significant representation in the data are copula *be* attested from file 10 (2;4.29) onwards with some consistency (see Table I section 7.2), and aspectual auxiliary *be* frome file 11 (2;7.8) onwards (see Table II, section 7.3). While for the copula there are a number of obligatory contexts even in the earliest files, for auxiliary *be,* obligatory contexts start to appear in file 8, (2;2.24) when the first examples of *-ing* forms are attested:

(14) File 8

*CAR: what doing?

The early presence of obligatory contexts for the copula is related to the presence of a number of naming constructions of the type *that (is) (a) x* that have been discussed in detail in section 9.6.3. By contrast, the appearance of aspectual auxiliary *be* is tied to the emergence of progressive forms of eventive verbs. For aspectual

auxiliary *be*, as was already observed for the copula, the only two contexts that have any significant representation are 3 p.s. and 3 p.p. contexts, 79 out of 82 forms are either 3 p.s. or 3 p.p.. Moreover, two *-ing* forms , *doing* and *going*, account for 63.41% of all verbs that appear with aspectual *be*. The remaining 15 verb types that appear with aspectual *be* account for less than 40% of present progressive occurrences.

Although there seems to be some indication that C. is beginning to realise the obligatoriness of aspectual *be* in progressive constructions around 2;7.8, the extremely limited knowledge of the morphological paradigm, together with the degree of lexical specificity of verb types that appear in progressive constructions, suggest that it would be premature to conclude that a productive AGR category underlies the production of such forms.

Further evidence that AGR is not active in C.'s English is given by the absence of auxiliary *do* in questions, and by its stereotyped use in negatives where it is found almost exclusively with *know* in the frozen phrase *I don't know*. Agreement and Tense contrasts are also virtually non-existent on lexical verbs. The proportion of verbs inflected for Number in present tense contexts never reaches a significant level (see Table VI, section 7.6). Moreover, there are only two verbs that are ever suffixed with Number marker *-s*: *come* and *go*.

As for Tense contrasts, the preterit forms in the corpus are not very reliable indicators that past tense forms are being paradigmatically contrasted with present tense forms. Of the 8 verb types that appear in the preterit (*fell, stopped, went, found, lost, said, wanted,* and *gave*), only 2, *want* and *stop*, are regular verbs that take the *-ed* suffix, the rest are suppletive irregular forms which may not have been learnt as forms contrasting with a corresponding present tense form at all. The number of present perfect forms is also extremely limited (N = 14), and the AGR/T marking auxiliary is present in only 4 occasions, less than 30% of the time. With such a small data set it is difficult to draw any conclusions as to the status of Tense and Aspect contrasts in present perfect forms, however the little evidence that is there does not suggest mastery of this verb form.

Despite the lack of morphosyntactic marking of Agreement and Tense that we have reviewed so far, C. uses a substantial number of modals (N = 72), elements that are typically assumed to be generated under T because of their defective agreement paradigm. C. seems to treat modals as a separate class from lexical verbs: they undergo

inversion in yes/no questions, they may take a negative clitic, and their verbal complements are always bare verbs, while verbal complements of lexical verbs are introduced by infinitival *to*:

(15) File 18

*CAR: all the giraffes can go.

(16) File15

*CAR: I want to make a garden.

However, although C. seems to have figured out the distributional properties of modals, his knowledge of this verb class can be said to coincide with knowledge of one single lexical item: *can*. One single modal, *can*, appears in 88.88% of all utterances containing a modal verb. In addition, two verbs alone, *put* (N= 25) and *go* (N= 16), account for more than 50% of all modal verbal complements (see Table IV, section 7.4).

In English C. displays the behaviour of a conservative learner, his principal strategy is to rely on a relatively small number of positional patterns, such as the semiproductive constructions *that's a x, it's a x,* and a limited verb lexicon whose items are combined with a small and highly predictable variety of subjects and complements (see for example the correlation between the modal *can* and the complements *put* and *go*). Another aspect of C.'s conservativism in the use of verbs is that he only ever commits omission errors. The conclusion is that there is no evidence that the child is aware of the obligatoriness of morphosyntactic Agreement and Tense marker by 3;1.25, hence the AGR and T category that are required for the checking of these formal features must be absent at this stage of the child's grammar. If this is the situation for English, what is the status of AGR and T in C.'s Italian? How does his use of verbs in Italian compare with English? The next section presents a summary of the conclusions reached in chapter 8 on C.'s use of verbs and contrasts it with the findings for English.

9.8.2. Evidence for AGR and T in Italian verbs

As previously reported in a number of studies on the acquisition of Italian morphology, like his monolingual peers C. produces a number of inflected forms as early as 1;10.27. The overall error rate for inflected forms in the present is also

consistent with findings in previous studies: 5.13%, oscillating between 0% for 1 p.p. and 13.88 for 3 p.p. (see Table XIII, section 8.3.3). Table XIV, section 8.3.3, reports the number of verb tokens and verb types across the 20 files for infinitive, imperative, present, present progressive, imperfect, and present perfect. Until file 10 (2;2.17), there is only one verb form per verb type per session, while from file 10 onwards the number of verb forms becomes increasingly larger than the number of verb types. This finding could be interpreted in one of two ways, either the trend is for a small number of verb types to appear with a large number of verb forms, or, alternatively, it could be the case that there is a relatively large number of verb types appearing with only one or two different inflections. It is clearly crucial to determine which pattern is applicable to C.'s acquisitional strategy; in the former case it looks like the child is acquiring new inflections on a verb-specific basis, he builds his knowledge of the paradigm around particular verbs, and the increasingly larger number of verb forms is given by the learning of different inflections for a small number of different verb types (*verb-based* acquisition). In the other scenario, where the number of verb forms is distributed across a wider verb type base, there would be stronger evidence for *paradigm-based* acquisition, where the child shows the ability to apply newly-acquired inflections to a large number of verb types, rather than focussing only on a handful of verb types.

The verb-by-verb table in Appendix 3, allows us to establish which of these two strategies C. has adopted. It does seem as if C. relies more on a verb-based strategy rather than a paradigm-based strategy. From file 13 (2;5.6) onwards, when the gap between the number of verb forms and verb types starts to widen, there is clear evidence that in each file the increasing number of verb forms is given by multiple inflections for only two or three different verb types. For each file from 13 to 20, Table I below summarises the distribution of the number of verb types that appear with one verb form (second column), two verb forms (third column), or more than two verb forms (fourth column), and figure 6 plots the relative proportions.

File	1	2	>2
13	16	2 andare, to go guardare, to look	3 cadere, to fall fare, to do/to make giocare, to play
14	15	6 andare, to go bere, to drink camminare, to walk leggere, to read portare, to bring prendere, to take	3 fare, to do/to make giocare, to play mangiare, to eat
15	10	10 ballare, to dance cadere, to fall dovere, must entrare, to enter girare, to turn mangiare, to eat potere, can soffiare, to blow stare, to stay volere, to want	3 fare, to do/to make mettere, to put spingere, to push
16	12	1 chiudere, to close	2 fare, to do/to make mettere, to put
17	16	3 giocare, to play guardare, to look volere, to want	2 fare, to do/to make prendere, to take
18	15	6 andare, to go cadere, to fall dovere, must guardare, to look mangiare, to eat mettere, to put	2 dare, to give fare, to do/to make
19	10	3 aiutare, to help perdere, to lose prendere, to take	8 andare, to go dire, to say dovere, must fare, to do/to make mangiare, to eat mettere, to put potere, can volere, to want
20	24	2 spaventare, to frighten volere, to want	8 andare, to go avere, to have dire, to say dovere, must fare, to do/to make mettere, to put prendere, to take trovare, to find

Table I. Distribution of Italian verbs with one, two, or more than two inflections per session

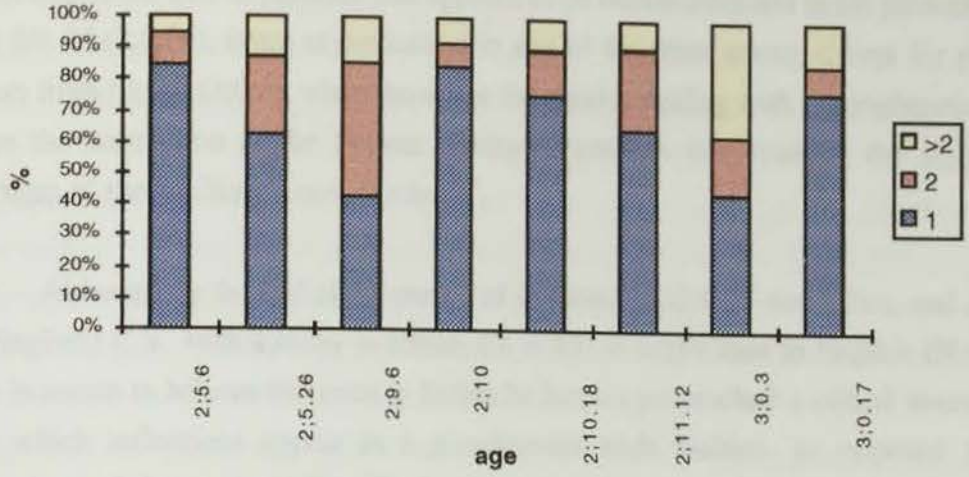


Figure 6. Proportion of Italian verbs with one, two, or more than two inflections per session

Except for file 19 (3;0.3) where the proportion of verb types occurring with more than two different verb forms reaches 37%, the percentage is on average around 10%. An average of 65% of verbs occurs with only one verb form, while the remaining 25% occur with two different forms. Therefore, even at a time where contrastive use of different verb forms emerges, it seems to be restricted to a small percentage of verbs. What is even more interesting is that it is the same verbs that appear session after session with more than one form. In particular the verb *fare, "to do/to make",* is present throughout the sessions from 13 to 20 with more than two forms. Altogether, only 14 verb types (17%) are used contrastively with more than two verb forms per session throughout the data collection period (*andare, "to go", cadere, "to fall", dare, "to give", dire, "to say", dovere, "must", fare, "to do/to make", giocare, "to play", mangiare, "to eat", mettere, "to put", potere, "can", prendere, "to take", spingere, "to push", trovare, "to find", volere, "to want"*). In fact until file 19 (3;0.3), there are only 8 verb types that appear with more than two forms per session; in file 19 (3;0.3) and in file 20 (3;0.17) another 6 verb types appear with multiple forms, thus almost doubling the previous figure.

The evidence presented above provides strong indications that C. relies heavily on a verb-specific strategy, new forms are learnt for a small number of verbs and it is the same verbs that appear with multiple forms. Although there is evidence of an emergent productive and contrastive use of verbs as from file 13 (2;5.6), there are also clear indications that contrasts are not found across the board for a large number of

verbs. In addition, not only are emerging contrasts verb-specific, there is also little evidence that contrasts that are established in one domain carry over to other domains. For example, the Person contrast that appears to be well established in the present tense from file 14 (2;5.26), is not as productive in any of the other tenses, except for present perfect from file 19 (3;0.3), where however the child is dealing with a periphrastic form where the acquisition of the Person contrast translates into learning the suppletive paradigm of the auxiliary *avere, "to have"*.

Although by the end of the period of observation (3;0.17 for Italian, and 3;1.25 for English) C.'s verb lexicon in Italian (N = 82) is larger than in English (N = 52), there is reason to believe that even in Italian he has not yet reached a critical mass point after which inflections appear in a paradigmatic-wide fashion, as opposed to the occurrence of the same verbs with the same inflections, and a gradual verb-specific build up of grammatical contrasts. Despite the difference in the overall size of verb lexicons in the two languages, the acquisition trend is similar, as shown in Figure 7 below:

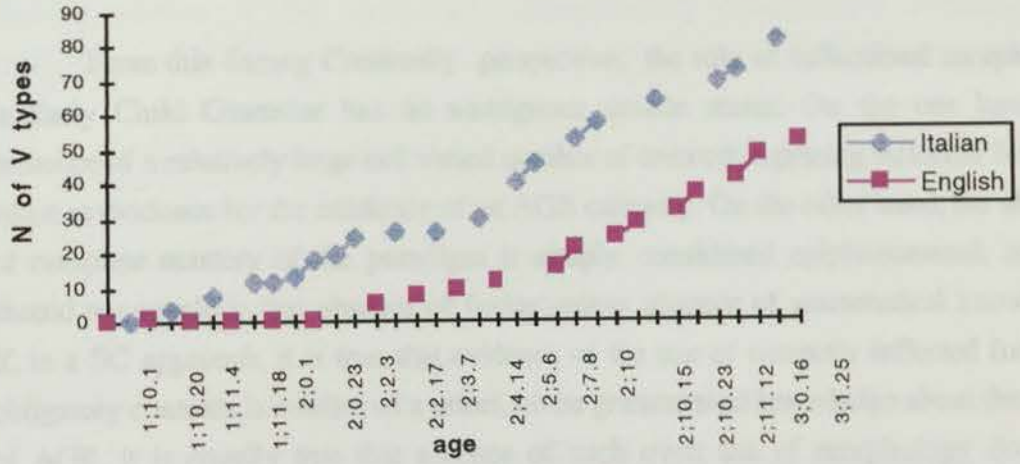


Figure 7. Verb vocabulary size in English and Italian

Both in English and Italian, the size of the verb lexicon increases gradually over time, but while in Italian there are a number of verbs from the earliest recordings, in English the first verbs emerge around 2;2.17

9.9. Theoretical implications

Given C.'s profile as a conservative verb-learner, what is the best characterisation of his grammatical knowledge? Crediting the child with an abstract notion of subject-verb agreement would be an overestimation of the child's grammar at this stage. There is simply not enough convincing evidence that the productive and contrastive use of verb forms goes beyond verb-specific learning. On the other hand, what kind of evidence would we need to infer that an abstract AGR category is present in the child's grammar? Those researchers that have argued for the early instantiation of an AGR category in early Italian child grammar (Hyams, 1986; Guasti, 1993/94), have relied on the number and variety of inflected forms, and on low or residual error rates in the earliest utterances of Italian-speaking children. They have dismissed the objections made by other reasearchers arguing for the incomplete mastery of the paradigm, by claiming that the fact that some inflections have not been acquired simply means that the *lexical learning* process is not completed yet (Hyams, 1992). This kind of argument rests on a qualitative distinction between grammar and the lexicon: grammatical notions are available a priori, as part of UG, lexical items need to be learnt on the basis of protracted exposure to the linguistic input.

From this Strong Continuity perspective, the role of inflectional morphology in Early Child Grammar has an ambiguous double status. On the one hand, the presence of a relatively large and varied number of correctly agreeing inflected forms is taken as evidence for the existence of an AGR category. On the other hand, the absence of complete mastery of the paradigm is simply considered epiphenomenal, and we should not conclude that absence of forms means absence of grammatical knowledge. If, in a SC approach, it is true that evidence of the use of correctly inflected forms in obligatory contexts is a reflex of a priori, given grammatical knowledge about the status of AGR, it is equally true that absence of such overt use of morphology does not undermine the theoretical position that a functional projection exists in the child's grammar.

Apart from the ambiguity inherent in this treatment of inflectional morphology, what is particularly problematic in this approach is the divide between lexical and grammatical acquisition. As mentioned in section 9.3, recent developments in a number of syntactic models, including generative models such as Chomsky's (1995) Minimalist Program, have reevaluated the role of the lexicon in syntax, giving it a much

more prominent role than it used to have in transformational grammar.[9] Moreover, a number of cross-sectional studies on the relationship between lexical and grammatical acquisition (Snyder, Bates & Bretherton, 1981; Bretherton, McNew, Snyder & Bates, 1983; Bates, Bretherton & Snyder, 1988) have also shown a tight link between vocabulary size and grammatical development. One of the findings from the Bates, Bretherton & Snyder (1988) study is that the best predictor of grammatical development at 28 months (in the heart of the "grammar burst"), as measured by MLU, is vocabulary size at 20 months (in the heart of the "vocabulary spurt"). These results highlight a crucial link between lexical and grammatical acquisition, where the latter cannot be dissociated from the former. In normally developing children, vocabulary size is highly correlated with grammatical development, and although grammar may fall behind vocabulary size, there are no reported cases of children that can stage productive grammar with a vocabulary of 100-300 words.[10]

In a large-scale study of the lexical and grammatical abilities of American English-speaking children between 16 and 30 months, Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly (1994) found a strong correlation between grammatical complexity, as measured by the 37-item sentence complexity scale on the Toddler version of the MacArthur Communicative Development Inventory, and productive vocabulary size. These findings for English have also been replicated by a number of cross-sectional studies on the acquisition of Italian between 16 and 30 months of age. Caselli, bates, Casadio, Fenson, Fenson, Sander & Weir (1995), Caselli, Casadio & Bates (1999) have found a similar function linking grammar and vocabulary size in Italian and English acquisition, despite striking differences between the two languages in the content of vocabulary and grammar.

The powerful relationship between lexical and grammatical growth has also been confirmed by a longitudinal study of 27 English-speaking children between 8 and 30 months of age (Goodman, 1995; Jahn-Samilo, 1995; Thal, Bates, Goodman & Jahn-Samilo, in press). Bates & Goodman (1997) note that, even when very different

---

[9]Lexicalist approaches such as Lexical Functional Grammar, Generalised Phrase Structure Grammar, Head-driven Phrase Structure Grammar, Categorial Grammar have been significant since the late '70s.

[10]An exception to the close association observed between lexical abilities and grammatical abilities is found in individuals affected by Down Syndrome where their lexical skills by far outweigh their grammatical skills (Thal, Bates & Bellugi, 1989; Mervis & Bertrand, 1993; Singer et al., in press). The argument that this impaired population is an example of the dissociation between separate lexical and grammatical modules (Pinker, 1991), has recently been challenged by Bates & Goodman (1997). Their alternative proposal is that this apparent dissociation may be caused by a deficit in aspects of information processing, such as poor auditory short-term memory, that are only indirectly related to grammar.

learner's profiles are compared, the relationship between vocabulary size and grammatical development still holds. In a comparison of one English-speaking late talker (272 words at 30 months), and one English-speaking early talker (315 words at 17 months) from the Thal et al.'s (in press) study, it is clear that both children are making progress in grammar that is directly proportional to their lexical abilities, although the two profiles are very different both in terms of vocabulary size and overall grammatical complexity. Marchman & Bates (1994) also report a powerful non-linear correlation between verb lexicon and the emergence of overregularisation of past tense forms, thus highlighting a particular type of lexical dependence between a specific class of lexical items and the development of a subcomponent of inflectional morphology. The implication we want to stress here is that one is likely to find different lexical "critical masses" for different types of grammatical knowledge, and hence different "critical masses" for the same syntactic category across different languages, depending on the transparency of the information it encodes. By this rationale, we would expect that in languages where Person deixis is encoded by a rich and transparent agreement system, like that of Italian or Spanish and Catalan for example, the critical mass for the acquisition of agreement morphology will be at a lower threshold than in a language like English where cues to agreement morphology are highly irregular and paradigmatically unpredictable.[11]

In sum, if lexical acquisition plays such a crucial role in the acquisition of morphology and syntax, it is surely not legitimate to treat the two as separate, qualitatively different mechanisms in language acquisition. The approach adopted here relies on lexical acquisition as a measure of grammatical development, hence until there is evidence that lexical items are being used by C. in a productive and contrastive fashion, we cannot speak of productive use of grammatical contrasts. Taking this argument a step further, we have to conclude that although, prima facie, C.'s knowledge of verbs is more sophisticated in Italian than in English, as confirmed by the number and variety of inflected forms in his corpus, this apparent advantage obscures the fact that his grammatical knowledge is still extremely lexically dependent. Even in Italian there is no clear evidence that he has indeed reached the critical mass point after which the verb base has become sufficiently large for the child to start making paradigm-based generalisations.

[11]Cann & Tait (1991) point out that while the acquisition of regular past tense suffixation in English requires the child to identify a binary paradigmatic distinction +/-Past, the acquisition of the 3 p.s. present form is more problematic because the child has to recognise the existence of a paradigmatic irregularity with a suffix that is otherwise morphophonologically regular.

In terms of the crosslinguistic emergence of AGR, we can conclude that neither in Italian nor in English does C. show productive knowledge of the abstract notion of subject-verb agreement, although in Italian where Person and Number contrasts are salient and transparent, the child has lexical knowledge of a larger number and variety of inflected forms. As for T, the very low number of preterit and present perfect forms in English prevents us from drawing any definitive conclusions other than that Tense/Aspect contrasts are just only beginning to be sporadically marked by the start of the third year. In Italian, a total of 30 verb types are found in the present perfect, but only a third of them (N= 11) appear with one or more Person/Number contrasts (*andare, "to go", cadere, "to fall", dare, "to give", dire, "to say", fare, "to do/to make", finire, "to finish", giocare, "to play", mettere, "to put", perdere, "to lose", prendere, "to take", trovare, "to find"*). The proportion of auxiliaryless present perfect forms is low overall (9.3%), an indication that, although the majority of verbs are accomplishment and achievement verbs, thus confirming the salience of inherent lexical aspect in the emergence of present perfect forms, nevertheless Tense and Person/Number features are made visible by the presence of the inflected auxiliary. By the end of the recording period (3;0.17) we also witness the first instances of imperfect tense, and thus C.'s first attempt to signal a grammatical Aspect contrast in the past between progressive imperfect and perfective present perfect. Of the three productive imperfect 3 p.s. verb forms in file 20 (3;0.17), two (*andava, "he/she went"*, and *faceva, "he/she did/made"*) also contrast with the corresponding 3 p.s. present perfect forms, showing an emergent Tense/Aspect contrast that is not so intimately tied to the lexical aspect of the verb. The number of verbs is however still very low to determine to what extent this emerging contrastive use of verb forms can abstract away from knowledge of specific lexical items.

9.10. Comparing like with like

9.10.1. The marking of Person deixis and the status of agreement morphology

In the three preceding sections we have proceeded to compare and contrast C.'s performance on inflectional morphology in the attempt to establish if, when, and how productive and contrastive use of morphosyntactic Agreement and Tense markers signals the acquisition of formal grammatical contrasts. This is standard practice not only in crosslinguistic studies of language acquisition, but also in studies in bilingual first language acquisition that focus on the acquisition of morphosyntax. Yet, when one compares languages that differ so significantly in the extent to which Person deixis is

marked on the verbal paradigm such as English and Italian, one cannot help but wonder whether one is truly comparing like with like. In other words, to what extent is it legitimate to compare the acquisition of inflectional morphology between English and Italian as a marker of Person deixis?

We know from a number of crosslinguistic psycholinguistic experiments (MacWhinney, Bates & Kliegel, 1984; MacWhinney & Bates, 1989) , that adult native speakers of English and Italian rely on different cues to identify the subject of a clause. In English the preferred strategy is to rely on word order, given the language's strict adherence to the SVO order. In Italian, still an SVO language, but with a freer word order and the possibility to have null or postverbal subjects, adult native speakers tend to focus primarily on verb morphology to identify the subject of a clause. The different weight that speakers of these two SVO languages give to inflectional morphology in parsing a sentence has been associated with the availability and the reliability of the cue provided by verbal inflection. The availability of a cue is measured as the ratio of cases in which the cue is available over the total cases in the task domain (MacWhinney & Bates, 1989). In the case of agreement inflections as markers of Person deixis in Italian (the task domain), the cue is highly available since all person/number combinations require that the verb take an inflectional ending. Similarly to availability, the reliability of a cue is determined as the proportion of cases in which a cue leads to the the correct assignment over the cases in which it is available. In Italian, the verbal inflection cue is not only 100% available, but it is also 100% reliable in that each person/number combination is uniquely identified by a different inflection.[12] In English, the cue of agreement with the verb is only available when there is a competition between two nouns and when those two nouns differ in number as in *The boys are chasing the girl.* By contrast, unlike in Italian where the distribution of subjects is much freer, the preverbal position cue in English is an excellent cue to the assignment of a noun phrase as the subject.

These crosslinguistically different parsing strategies that adult native speakers use reflect the validity of verbal agreement and preverbal position as possible cues to identify noun phrases as subjects. Because the reliance on one or the other strategy to parse the linguistic input depends on statistical distributional properties of the different languages, there is reason to believe that children might indeed exploit the same type of

[12]The higher cue validity of verbal agreement in Italian does not exclude that adult native speakers will also use the preverbal position as predictor of subjecthood, in fact the two cues can be used together to optimise parsing. It is only when the two cues are in conflict that verbal agreement will win out over preverbal position.

information as adults to parse the incoming input. If this is along the right lines, one would indeed expect Italian-speaking children to focus on the acquisition of verbal morphology, while English-speaking children should focus more on the production of overt subjects to mark person deixis. Although, as far as we know, these claims have not specifically been addressed simultaneously before, a number of independent studies in the literature give credit to this hypothesis.

As has already been illustrated in some detail in the previous sections of this chapter, there is evidence that monolingual English-speaking children, as well as bilingual English-speaking children, take a considerable length of time before they make productive use of agreement Number morphology (see Cazden, 1968; Brown, 1973; Radford, 1990; Paradis & Genesee, 1996, 1997; Sinka & Schelleter, 1998; Wexler, 1994; 1999, among many others). Italian children, by contrast, use a relatively large number of inflected forms from their earliest utterances, although productive contrasts take some time to emerge (see Hyams, 1986, 1992; Pizzuto & Caselli, 1992, 1994; Guasti, 1993/94). It seems that by learning and producing a number of inflected forms, Italian-speaking children are focusing on the right type of cue to mark Person deixis, and that, at the same time, English-speaking children are temporarily neglecting information that is both not readily available, nor highly reliable on the verbal paradigm. Nevertheless, the marking of Person deixis is as essential in English as it is in Italian, although the two languages differ to the extent in which this information is encoded (verbal morphology in Italian vs. overt subjects in English).

At the same time as there is evidence that Italian children are successfully beginning to mark Person contrasts on verbs, there are also indications that English-speaking children are sensitive to the cue expressed by overt subjects in their language. The very extensive literature on null subjects in English, and in a number of other non-null-subject language, has tended to focus on the non-target use of null subjects in the initial phases of acquisition. A number of explanations have been offered for this null-subject phenomenon in early child language: performance limitation accounts (L. Bloom, 1970; P. Bloom, 1990, 1993; Valian, 1990, 1991, Valian & Eisenberg, 1996), parameter missetting accounts (Hyams, 1986), topic-drop accounts (Hyams & Wexler, 1993; Haegeman, 1994), missing functional projections accounts (Rizzi, 1994a, 1994b), underspecification of Number accounts (Sano & Hyams, 1994; Hoekstra & Hyams, 1995; Hyams, 1996, Hoekstra, Hyams & Becker, 1997), Optional Infinitive accounts (Schütze & Wexler, 1997; Wexler, 1994, 1999). What all of the above competence-limitation approaches have in common is a view of English Early Child

Grammar as deficient in one way or another: it may present missing or optional functional projections, underspecification of formal features, or missetting of a parameter, but all of these accounts view English children's production of null subjects as a sign that their grammatical abilities are immature.

A different perspective on early subjects in early child grammar is provided by Valian (1991) and Valian & Eisenberg (1996) in two studies on subject use in American English, Italian and Portuguese. The emphasis in these two studies is on the extent to which English-speaking children actually *use* overt subjects, as opposed to the fact that they also *omit* them some of the time. The fact that English-speaking children's performance is in fact significantly different from that of children acquiring null-subject languages has often been neglected in favour of the exclusive treatment of null subjects. In a comparative study of subject use in a group of American English-speaking children (age range between 2;0 and 2;5), and a group of Italian-speaking children (age range between 2;0 and 2;6), Valian (1991) found that there was a significant difference between the two language groups. American children with an MLU below 2.0 morphemes used subjects an average of 69% of the time, while their Italian counterparts, matched for MLU, only produced subjects 30% of the time. Similar findings were more recently replicated by Valian & Eisenberg (1996) in a study on subject use in children acquiring Brazilian and European Portuguese. In all of the three MLU-matched groups the Portuguese-speaking children and English-speaking children were divided into, subject use was dramatically different in the two languages. In group I (MLU lower than 2) and group II (MLU between 2 and 2.5) the American children used subjects more than twice as often as the Portuguese-speaking children, and in group III (MLU between 2.5 and 4) more than 1.5 times as often.

This statistically significant difference in the use of subjects between children acquiring a poorly inflected non-null-subject language like English, and children acquiring rich agreement null-subject languages like Italian and Portuguese must surely reflect the early sensitivity of the former to the requirement that subjects be expressed overtly in their language. In addition, as noted earlier, subjects in English are Person deixis markers and encode some of the relevant information that is expressed by verbal morphology in Italian. The suggestion being made here is that English-speaking children pay attention to subjects because they are sensitive to their statistical frequency, and because they are the most reliable cue they have to express Person deixis in their language. The fact that English-speaking children produce subjects significantly more often than their peers acquiring rich-agreement null-subject

languages goes some way to support the hypothesis put forward above whereby Person deixis will be expressed by the use of overt subjects in English, and by use of verbal morphology in Italian.

9.10.2. C.'s differential use of subjects in English and in Italian

In section 7.8 we noted how, despite the fact that C.'s use of verbs does not show productive evidence of Agreement and Tense marking, nevertheless from file 9 (2;4.7) onwards, subjects are supplied an average of 85% of the time. A breakdown of subject use per verb form in Table X, section 7.8, also reveals that subjects are present an average of 79.41% of the time with bare forms, an unexpected result for researchers arguing for a correlation between null subjects and Root Infinitives (RIs) (see Hoekstra, Hyams & Becker, 1997 and references cited therein). The prediction that RIs should cooccur with null PRO subjects or indeterminate DPs, i.e. bare singular nouns, stems from the hypothesis that the deficit underlying both RIs and null subjects/indeterminate DPs is underspecification of the Number feature. Only bare singular nouns are expected with bare verbal forms with no overt Number specification.

In C.'s data, out of a total of 54 overt subjects of RIs, 24 are instances of demonstrative pronouns *this* and *that*, 2 are bare nouns, 7 are proper names, 14 are personal pronouns (*I, he/she, we*), and 7 are DPs including a determiner, one of which contains a plural noun (*the squirrels*). The two demonstrative pronouns are all-purpose placeholders with no referential specification; together with the 2 bare nouns they make up for 48.14 of all overt subjects found with RIs. The remaining 51.86% of RIs' overt subjects are personal pronouns (25.92%), proper names (12.96%), and determinate DPs (12.96%). The proportion of [+specific] overt subjects with RIs is thus far from negligeable, it is a phenomenon that calls for an explanation.

In section 7.8 we also noted that the number of subjects increases as a joint function of increasing MLU and increasing verb use, although it is not positively related to the increase in the use of inflected form. The solution proposed to this conundrum was to see C.'s use of subjects as an emerging realisation independent of the acquisition of Agreement and Tense marking. Ingham (1992: 148-149, my emphasis) has made a similar proposal for the dissociation between inflectional elements and subject use in a monolingual English child between 2;5 and 3;0:

> It seems appropriate to suggest that the disappearance of the Optional Subject Stage may not necessarily be prompted by changes occurring elsewhere in children's developing grammars, but that children may attend to the obligatoriness of subjects in English as *an independent grammatical property acquired in its own right.*

We further qualify this "independent grammatical property" as the need to mark Person deixis, the most reliable way in English being the use of an overt subject. In Italian, the morphological marking of Agreement and the expression of person deixis coincide since the language is a Person-marking language; in English on the other hand, where only Number, and not Person is marked by inflectional morphology, Person is made visible only by the use of an explicit subject. If this hypothesis is correct we should find that C. uses subjects in English significantly more often than he does in Italian. Figure 8 below plots the percentages of overt subjects in Italian and English over a period of 15 months.

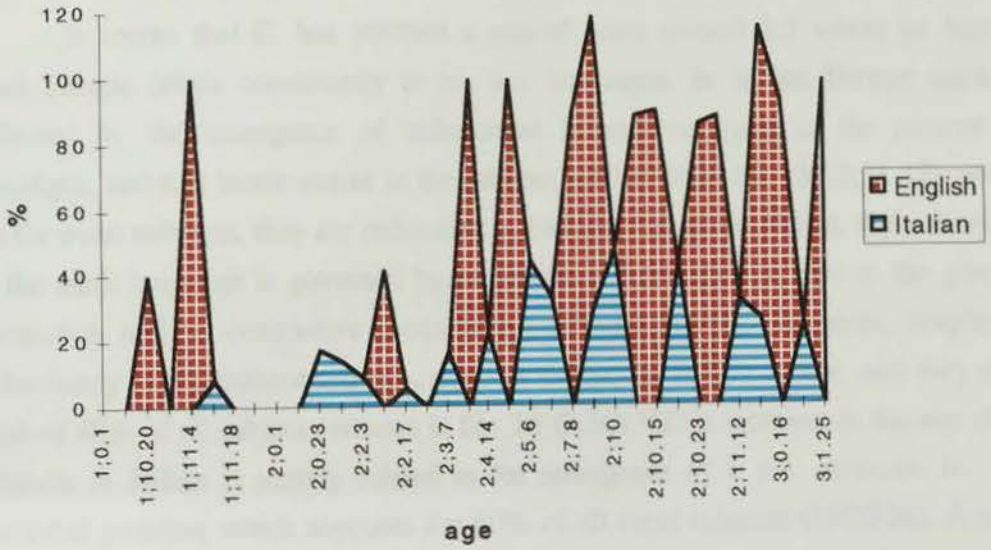


Figure 8. Proportion of verbal predicates with an overt subject in English and Italian

Figure 8 shows that indeed there is an overall higher proportion of subjects in English throughout the period of observation. The average percentage for English is around 85%, while in Italian it is around 22%, an even larger difference than that observed by Valian between her American and Italian subjects (Valian, 1991), and American and Portuguese-speaking children (Valian & Eisenberg, 1996). There is

evidence that, like his monolingual English-speaking peers in Valian's (1991) study, C. is sensitive to the requirement that English subjects must be expressed overtly, although none of the morphological correlates of Agreement are in place yet. The obvious crosslinguistic difference observed in the use of subjects in English and Italian also shows that the child is not applying an overt subject strategy across the two languages. He is clearly paying attention to the different requirements of the two languages he is exposed to, and he treats them as separate, self-contained problem spaces.

A further piece of evidence that supports the hypothesis that C. is focusing on different morphosyntactic elements in the two languages to express the same notion of Person deixis, is given by the chronological parallel between the emergence of a substantial proportion of subjects in obligatory contexts in English, and the emergence of contrastive use of inflectional morphology in Italian. As shown in Figure 8, subjects in English reach a significant percentage in obligatory contexts in file 9 (2;4.7) and 10 (2;4.29), and in Italian it is around the same time, file 13 (2;5.6), that the first productive Person contrasts begin to emerge (see Table XVII, section 8.3.4).

It seems that C. has reached a critical point around 2;5 where he begins to mark Person deixis consistently in his two languages. In Italian, Person marking is reflected by the emergence of inflectional Person contrasts in the present tense paradigm, and to a lesser extent in the present perfect from file 18 (2;11.12) onwards. As for overt subjects, they are redundant as Person markers in Italian, their distribution in the adult language is governed by pragmatic requirements related to the given/new distinction and to contrastive focus (see Calabrese, 1991). However, despite their redundancy as Agreement markers, subjects do appear in C.'s Italian, and they reach a peak of 48% of all subject contexts in file 16 (2;10).[13]This increase in the use of overt subjects in Italian is mainly related to the emergence of 1 p.s. pronoun *io, "I"*, in preverbal position which accounts for 60% of all overt subjects (197/326). A number of these occurrences of 1 p.s. pronoun *io* seem to be related to the marking of contrastive focus, although in adult Italian the only subject position for contrastive focus is the postverbal and not the preverbal one:

[13]Because Italian is a language where subjects are not obligatorily expressed, we cannot speak of obligatory syntactic contexts for subjects to the same extent as this is appropriate in English where overt subjects are required for all finite verbs. We can however note that subjects are possible with a finite verb, although they may not always be pragmatically appropriate.

(17) File 9

*LUD: allora chi va a cercare Orsetto?
%eng: so who is going to look for Teddy?
*CAR: io lo cerco.
%eng: I look for it.

Although C.'s reply in (17) is grammatical, it is pragmatically inadequate; an adult Italian speaker would have said *Lo cerco io, "Look for it I"*, with the subject in contrastive focus in postverbal position.

Because C. is a English-Italian bilingual child one might suggest that this inappropriate focus strategy could be a transfer from English, where no postverbal position is possible for contrastive focus and instead a prosodic strategy is available whereby the subject is stressed as in ***I*** *do it*. There are however at least two reasons why transfer from English is not the best explanation for C.'s use of 1 p.s. pronoun preverbal subjects in Italian. Firstly, there is no stress on the preverbal pronoun, as one would in fact expect if the child had transferred the prosodic focussing strategy from English. Secondly, in a longitudinal study of a monolingual Italian-speaking child, Camilla, Antelmi (1997) reports that Camilla went through a phase around 22-23 months, in which *io, "I"*, becomes productive and its presence seems almost obligatory. In fact, until 24 months null subjects account for just over 30% of subject contexts, 40% are represented by preverbal subjects, and the remaining proportion by postverbal subjects.

If we discard the transfer account, an alternative explanation for C.'s pragmatically inadequate use of preverbal subjects could simply be that contrastive focus is initially marked by the presence of a subject vs. its absence. In other words, focus is not marked by the contrastive placement in postverbal vs. preverbal position, but by overt vs. null subject.

In addition to preverbal subjects that seem to mark contrastive focus as exemplified in (17) above, there are also a number of subjects in C.'s Italian that are more like markers of agentivity. In this respect C. fits the profile of an ego-anchored child who primarily refers to himself as the main participant when he uses an overt subject.[14] There are in fact only a handful of pronouns other than *io*: 9 tokens of 2 p.s.

[14]The term ego-anchored is not to be confused with Piaget's notion of egocentric speech (Piaget, 1926). Instead, the term ego-anchored simply implies that the child's linguistic references to main participants tend to refer to self. See Budwig (1995) for an analysis of agentivity, control and the distributional patterns of subject referents.

tu, "you", and one example of 3 p.p. *loro, "they"*, throughout the Italian corpus. In sum, when C. does use 1 p.s. pronouns, they are either focused subjects, or ego-anchored markers of agentivity. However, the single most important fact about C.'s use of subjects in Italian is that they occur at significantly lower rates than in English, thus signalling a crosslinguistic difference in the weight the child assigns to overt subjects as markers of Person deixis in the two languages.

In conclusion, there is evidence for the hypothesis that C. starts to mark Person deixis at the same time in both English and Italian. Because the two languages differ as to where and how they encode this type of information, as expected, the child focuses on those cues that are most available and reliable in each language: verbal morphology in Italian and overt subjects in English. While a comparison of the emergence of morphosyntactic Agreement and Tense showed an advantage in Italian, if not in terms of productive contrasts at least in the sheer number of inflected forms; a comparison of the marking of Person deixis dissociated from control of the Agreement paradigm has revealed a developmental parallel between the two languages. The ability to treat morphosyntactic cues as language-specific shows that the child is indeed sensitive to the type of statistical information available in the input, and that he can treat the two languages as separate, independent grammatical spaces.

9.11. Summary

This comparative review of the emergence of the nominal and the verbal systems in C.'s English and Italian has revealed a number of differences and similarities. Starting from the fourth research questions in 9.1. on the Separate Development Hypothesis, we can conclude here that there is evidence that the child is indeed acquiring the two languages as independent and separate grammatical systems. There are no instances of morphological mixing, C. never suffixes Italian inflectional morphology to English stems or viceversa. The few instances of mixing are unidirectional from Italian to English, and they are limited to lexical borrowing of nouns that have no translation equivalents in English in the child's vocabulary:

(18) File 10

*CAR: it's a camion@s [*].
%err: camion = lorry $LEX $LAN[15]

[15]The $LEX code indicates that it is an error at the lexical level, and the $LAN code indicates a wrong language choice.

As for syntactic mixing there are no indications that C. has transferred the grammatical feature values of one language to the other, although neither the nominal system nor the verbal system have reached a level of complexity in the two languages that would allow us to disconfirm this claim. In the nominal system, English and Italian differ with respect to the distribution of definite articles with plural and mass generics, English disallows them while they are required in Italian (cf section 5.3):

(19)a I love (*the) wine.
b Mi piace *(il) vino.

(20)a I hate (*the) spiders.
b Odio *(i) ragni.

As reported in sections 6.2 and 6.4, the number of plural and mass generics is not very high in either language, in particular plural generics are virtually non-existent. It must be noted however, that those mass nouns that do occur in English and Italian are used appropriately in both languages. In other words, we do not find non-target bare mass generics in Italian, or mass generics with an article in English:[16]

(21)a @I want the juice.
b @Voglio succo.

As for plural generics, there are a couple of examples in Italian where the definite article is missing, which might indicate a transfer from English. However, it must be noted that, even by the end of the period of observation, there is still a residual error rate whereby the absence of a determiner with an Italian plural generic could be the result of the optionality of articles rather than transfer from English.

As for the question of syntactic transfer in the verbal domain, it is difficult to interpret the data from these early stages of grammatical development. One could argue that transfer in the syntax of verbs would be supported, for example, by evidence that the child had assumed the strong Agreement and Tense features of Italian verbs for his English verbs thus resulting in overt raising before Spell Out. If that were the case, one might find utterances like (22) where the verb has raised to the left of negation to higher AGR and T projections to check off its features:

(22) @I want not this.

[16]This claim clearly applies to the stage in which articles are supplied in obligatory contexts at least 90% of the time. Before this point bare nouns account for a large proportion of C.'s nominals in both languages.

This kind of error is unattested in the linguistic development of monolingual English-speaking children, and should we find examples such as (22), we might reasonably infer that they are the result of transfer from Italian. These occurrences of an English verb to the left of negation are unattested in C.'s corpus and, as far as the literature reports, they are also unattested in other bilingual children acquiring English and another language where verbs have strong Agreement and Tense features (see Paradis & Genesee, 1997 for English-French bilinguals; DeHouwer, 1990, for an English-Dutch bilingual; Sinka & Schelleter, 1998 for two English-German and English-Latvian bilinguals). There are a number of reasons why we might in fact expect not to find utterances like (22) in an English-Italian bilingual context. Firstly, the Italian negative head *non* is a clitic that moves along with the verb to AGR to a preverbal position, resulting in a very similar word order to English Subj+Neg+Verb:

(23) Non voglio mangiare.
"(I) don't want to eat"

Even monolingual English-speaking children initially treat *don't* simply as a negative marker and not as the contracted form of *do not,* where *do* is treated as an auxiliary verb outside VP, and we have argued that the few occurrences of negated utterances in C.'s corpus give no clear indication that *do* is actually being used productively as an auxiliary. If this is the case, negated utterances of the type (Subj)+Neg+Verb are ambiguous, it is in fact the correct word order for both English and Italian, although the derivation of the same superficial word order implies treating *don't* and *non* as very different elements, at least in the adult grammar. Examples from negative utterances can therefore not be used as unambiguous evidence that the child is treating English and Italian as two different grammatical systems, although the data does not disconfirm this hypothesis. Furthermore, the absence of *do* in C.'s interrogatives, and question formation through rising intonation is a well-documented phase through which monolingual English-speaking children go through as well, and thus it cannot be taken as evidence of transfer from Italian (Klima & Bellugi, 1966; Radford, 1994, Stromswold, 1995).

Aside from lexical and morphological evidence for the separate development of two grammatical systems, is there any syntactic evidence at all that would further support the Separate Development Hypothesis? We believe that the data on the distribution of subjects in English and Italian in section 9.10.2 provides unambiguous evidence that C. is indeed dealing with his two languages in a separate and independent

way. The very different proportions of overt subjects in English (85% on average) and in Italian (22% on average) give credit to the hypothesis that the child is sensitive to the distribution of overt subjects in his two languages, although, as we observed, he has not yet entirely mastered the pragmatics that regulate the distribution of overt subjects in Italian.

Summing up the results of our findings on the acquisition of grammatical marking we conclude that, in both the nominal and the verbal domain, lexical learning shapes grammatical development. In both languages, although the effect is considerably more visible in English, C.'s linguistic production is kickstarted by a small number of limited scope formulae, and by a conservative use of nominal and verbal forms, the same nouns and verbs tend to appear again and again in the same constructions: the best predictor of a verb's use is the immediately preceding occurrence of the same verb as opposed to use of other verbs at the same time.[17]

Although there is a degree of similarity in the path to the emergence of grammatical marking and its functional realisations in C.'s two languages, there are also clear differences, the most obvious of which is the overall lead-lag pattern with Italian ahead of English by 2 to 5 months on average. The explanation we have invoked to account for such a discrepancy is a typological one, where the acquisition of grammatical contrasts is facilitated in Italian by the transparency, availability, and reliability of inflectional and free-standing morphology.

As a concluding remark, in addition to the typological explanation of crosslinguistic differences, we would also like to emphasise the difference between the two environments in which C. is exposed to Italian and English. While Italian is the language spoken to him by both his parents, his two brothers, and his babysitters (home language), he is directly addressed in English only at the nursery where he shares the time and attention of one, at times two, nursery staff with seven or eight other children (community language). A number of studies comparing dyadic and polyadic situations in adult-child interaction (Sylva, Roy & Painter, 1980; Schaffer & Liddel, 1984; Pellegrino & Scopesi, 1990) have emphasised that, not only does quantity of child-directed speech decrease when the adult-child ratio decreases, but also that the quality of the interaction changes significantly. In a group situation with more than one child, the adult is more likely to use peremptory and prohibitive language, a

[17]Tomasello (1992: 256) makes this same claim regarding the status of his daughter's early use of verbs.

substantial number of children's bids for attention are ignored, there is considerably less joint involvement with individual children, less conversation-eliciting behaviour, and the number of turns tends to be smaller than in dyadic situations (see section 3.6). Research on the effect of maternal conversational style on children's grammatical development has also shown that "for normally developing children at least, maternal conversational style contributes to the development of grammar above and beyond the contribution of the child's pragmatic capacities" (Rollins & Snow, 1998). If not only the quantity, but also the quality of the language a child is exposed has an effect on her grammatical development, it is reasonable to assume that the very different types of environments, and the different interactional styles of C.'s Italian-speaking interlocutors and his English-speaking caretakers may well have had an impact on his acquisition of the two languages.

In studies of bilingual first language acquisition the primary concern has often been to assess whether the child is exposed to the two languages on a regular, daily basis, i.e. whether the child has a fairly equal and balanced access to both languages. Although more sociolinguistically oriented case studies of bilingual first language acquisition have paid attention not only to the quantity (Lanza, 1992, 1997; Döpke, 1992), but also to the quality of the bilingual input, with a few exceptions (DeHouwer, 1990) studies focussing primarily on morphosyntactic development have tended to overlook the fundamental role of the input. We believe that the social environment in which the bilingual child acquires her two languages, and the nature of the interaction the child engages in with speakers of the different languages deserve more attention than they have had so far in more syntax-oriented research. Exposure to two languages in different environments with different people is yet another variable that may turn out to have a non-trivial explanatory power in studies of bilingual morphosyntactic development.

Chapter 10

Conclusions

10.1. Empirical findings

This case study of the morphosyntactic development of one English-Italian bilingual child has made a number of empirical and theoretical contributions which are of interest for researchers working on childhood bilingualism and language acquisition in general. This section deals with the main empirical findings, while section 10.2 illustrates the methodological and theoretical implications of this work, and section 10.3 makes suggestion for further research.

The first practical contribution of this piece of research is the availability of a new bilingual corpus of English-Italian child data, a language combination for which there were no available corpora to date. The data reported on here have not been made publicly available yet, but will be contributed to CHILDES as soon as they have been at least partially double checked by a second transcriber. Despite a growing number of new additions to CHILDES of child data in a number of languages, English is still the language for which there is the single highest number of corpora. Serious crosslinguistic research must be based on as large a body of data as possible, coming from a number of typologically different languages. For bilingual child data the collection of new corpora is even more necessary in the absence of a normative baseline of what can be considered as "typical" bilingual development.

As regards more specifically the empirical findings of this work, we have identified three areas in which our results contribute to current research in childhood bilingualism and language acquisition in general. Firstly, in line with previous studies on the morphosyntactic development of young bilingual children (Meisel, 1990; 1994; De Houwer, 1990; Paradis & Genesee, 1996, 1997; Sinka & Schelleter, 1998), the evidence presented in this work has confirmed that children who are exposed to two languages from birth in a separate fashion develop two independent, self-contained

grammatical systems. In C.'s acquisition of DPs, and in his acquisition of Agreement and Tense we have observed a language-specific developmental schedule, and the absence of any transfer of feature values from one language to the other. The strongest piece of evidence for C.'s separate and independent development of two grammatical systems is his marking of Person deixis (see section 9.10.2). In Italian, a Person-marking language, we can say that there is an emerging notion of Person by 2;5, when C. begins to use inflectional morphology in the present tense to mark person contrasts systematically. At the same time in English the proportion of subjects in obligatory contexts reaches 100%, while Agreement and Tense contrasts are still non-existent. This is clear evidence that the child must be focussing on different, language-specific properties of the input to solve the all-important communicative problem of marking Person deixis.

Secondly, we have identified language-specific strategies in C.'s acquisition of English and Italian morphosyntax, which indicate that the nature of the input the child is exposed to is crucial in shaping the emergence of Functional Categories. While in both languages lexical specificity plays a significant role, especially in the earliest stages of language production, there is an indication that in English he relies on lexically-specific patterns to a greater extent than he does in Italian. English multiword utterances tend to include a smaller number of verb types which do not contrast for Agreement and Tense, and word combinations tend to appear in highly predictable and stable patterns. See for example C.'s extremely limited use of verbal complements with modal verbs and specifically *can* (section 7.4).

Thirdly, we have shown how the contexts in which the child is socialised in his two languages play a fundamental role not only in the age and rate of acquisition of lexical and grammatical knowledge, but also in the way the child approaches the language learning task. The general trend is that of a lead-lag pattern with Italian an average of two to five months ahead of English in the emergence of a number of morphosyntactic devices (see section 9.3.1. for DPs, and section 9.8.1. for the emergence of Agreement and Tense). A possible explanation offered for this developmental asynchrony has been the different typology of the two languages. In Italian, cues to the acquisition of morphology are highly transparent, reliable and consistent, and it is only natural that the acquisition process should be facilitated in such a language. In English, by contrast, morphological cues to Agreement and Tense are less salient and even monolingual children have difficulties in extracting paradigmatic patterns in the absence of unambiguous evidence.

Bibliography

Aldridge, M. (1989). *The acquisition of INFL.* Bloomington: Indiana University Linguistics Club Publications.

Ambalu, D., Chiat, S. & Pring, T. (1997). When is it best to hear a verb? The effects of the timing and focus of verb models on children's learning of verbs. *Journal of Child Language, 24,* 25-34.

Andersen, R.W. (1993). Four operating principles and input distribution as explanation for underdeveloped and mature morphological systems. In K. Hyltenstam & A. Viborg (eds.), *Progression and regression in language*, pp. 309-339. Cambridge: Cambridge University Press.

Antelmi, D. (1997). *La prima grammatica dell'Italiano: indagine longitudinale sull'acquisizione della morfosintassi italiana.* Bologna: Il Mulino.

Antinucci, F. & Miller, R. (1976). How children talk about what happened. *Journal of Child Language, 9,* 627-643.

Bamberg, M.G.W. (1987). *The acquisiton of narratives: learning to use language.* Berlin: Mouton de Gruyter.

Bates, E., Camaioni, L., & Volterra, V. (1979). The acquisition of performatives prior to speech. In E. Ochs & B Schieffelin (eds.), *Developmental Pragmatics.* New York: Academic Press.

Bates, E., Bretherton, I. & Snyder, L. (1988). *From first words to grammar: individual differences and dissociable mechanisms.* New York: Academic Press.

Bates, E. & MacWhinney, B. (1989). *The crosslinguistic study of sentence processing.* Cambridge: Cambridge University Press.

Bates, E. & Goodman, J.C. (1997). On the inseparability of grammar and the lexicon: evidence from acquisition, aphasia, and real-time processing. *Language and Cognitive Processes, 12,* 507-586.

Bates, E. & Goodman, J.C. (1999). On the emergence of grammar from the lexicon. In B. MacWhinney (ed.), *The emergence of language.* Mahwah, NJ: Erlbau,.

Bennett-Kastor, T. (1983). Noun phrases and coherence in child narratives. *Journal of Child Language, 10,* 135-149.

Bergman, C. (1976). Interference vs. independent development in infant bilingualism. In G. Keller, R. Tescner and S. Viera (eds.), *Bilingualism in the bicentennial and beyond.* New York: Bilingual Press.

Berkele, G. (1983). *Die Entwicklung des Ausdrucks von Objektreferenz am Beispiel der Determinanten. Eine empirische Untersuchung zum Spracherwerb bilingualer Kinder (Französisch/Deutsch).* Master thesis, University of Hamburg.

den Besten (1989). *Studies in West Germanic syntax.* Amsterdam: Rodopi.

Bickerton, D, (1981). *Roots of language.* Ann Arbor, MI: Karoma.

____ (1984). The language bioprogram hypothesis. *Behavioural and Brain Sciences, 7,* 173-188.

Bloom, L. (1970). *Language Development.* Cambridge, Mass.: The Mit Press.

Bloom, L., Lifter, K. & Hafitz, J. (1980). Semantics of verbs and the development of verb inflection in child language. *Language, 56,* 386-412.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry, 21,* 491-504.

Bobaljik, J. (1997). If the head fits...: on the morphological determination of Germanic syntax. *Linguistics, 35,* 1029-1055.

Bobaljik, J. Jonas, D. (1996). Subject positions and the role of TP. *Linguistic Inquiry, 27,* 195-236.

Borer, H. & Rohrbacher, B. (1997). Features and projections: arguments for the full competence hypothesis. *Proceedings of BUCLD 22,* 24-35.

Boser, K., Lust, B., Santelmann, L. & Whitman, J. (1992). The syntax of V2 in early child German grammar: the strong continuity hypothesis. *Proceedings of NELS, 22,* 57-66. Amherst, Mass.: Graduate Linguistics Student Association.

Bottari, P., Cipriani, P. & Chilosi, A.M. (1993/94). Protosyntactic devices in the acquisition of Italian free morphology.*Language Acquisition, 3,* 327-369.

Braine, M. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development, 41* (Serial No. 164).

Brent, M.R. (1993). From grammar to lexicon: unsupervised learning of lexical syntax. *Computational Linguistics, 19,* 243-262.

Bresnan, J. (in press). *Lexical functional syntax.* Oxford: Blackwell.

Bretherton, I., Mcnew, S., Snyder, L. & Bates, E. (1983). Individual differences at 20 months. *Journal of Child Language, 10,* 293-320.

Brown, R. (1973). *A first language: the early stages.* London: George Allen and Unwin.

Brown, R. & Fraser, C. (1963). The acquisition of syntax. In C. Cofer & B. Musgrave (eds.), *Verbal behaviour and learning: problems and processes,* pp. 158-201. New York: McGraw-Hill.

Brown, R. & Bellugi, U. (1964). Three processes in the child's acquisition of syntax. *Harvard Educational Review, 34,* 133-151.

Bruner, J., Oliver, R. & Greenfield, P. (1966). *Studies in cognitive growth.* New York: Wiley.

Budwig, N. (1995). *A developmental-functional approach to child language.* Mahwah, NJ: Erlbaum.

Burzio, L. (1986). *Italian syntax.* Dordrecht: Reidel.

Calabrese, A. (1991). Some remarks on focus and logical structure in Italian. *Harvard Working Papers in Linguistics, 1,* 91-127.

Cann, R. & Tait, M.E. (1991). The persistence of principles. Unpublished manuscript, The University of Edinburgh.

Cartwright, T.A. & Brent, M.R. (1997). Syntactic categorization in early language acquisition: formalising the role of distributional analysis. *Cognition, 63,* 121-170.

Caselli, M.C., Leonard, L., Volterra, V. & Campagnoli, M.G. (1993). Towards mastery of Italian morphology: a cross-sectional study. *Journal of Child Language, 20,* 377-393.

Caselli, M.C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L. & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development, 10,* 159-199.

Caselli, M.C., Casadio, P. & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language, 26,* 69-111.

Cazden, C. (1968). The acquisition of noun and verb inflections. *Child Development, 39,* 433-448.

Chien, Y-C & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition, 1,* 225-295.

Chini, M. (1997). *Genere grammaticale e acquisizione: aspetti della morfologia nominale in Italiano L2.* Milano: Francoangeli.

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

____ (1995). *The Minimalist Program.* Cambridge, Mass.: The Mit Press.

Cipriani, P., Chilosi, A.M., Bottari, P. & Pfanner, L. (1993). *L'acquisizione della morfosintassi in Italiano.* Padova: Unipress.

Clahsen, H. (1990). Constraints on parameter setting: a grammatical analysis of some acquisition stages in German child language. *Language Acquisition, 1,* 361-391.

Clahsen, H. & Penke, M. (1992). The acquisition of agreement morphology and its syntactic consequences: new evidence on German child language from the Simone corpus. In J. Meisel (ed.) *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition.* Dordrecht: Kluwer.

Clahsen, H., Penke, M. & Parodi, T. (1993/94). Functional categories in early child German. *Language Acquisition, 3,* 395-429.

Clahsen, H., Eisenbeiss, S. & Vainikka, A. (1994). The seeds of structure: a syntactic analysis of the acquisition of Case marking. In T. Hoekstra & B. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar.* Amsterdam: John Benjamins.

Clahsen, H., Eisenbeiss, S. & Penke, M. (1996). Lexical learning in early syntactic development. In H. Clahsen (ed.), *Generative Perspectives on Language Acquisition,* pp 129-160. Amsterdam: John Benjamins.

Clark, E. (1985). The acquisition of Romance, with special reference to French. In D. Slobin (ed.), *The crosslinguistic study of language acquisition.* Vol i. *The Data.* Hillsdale, NJ: Erlbaum.

____ (1996). Early verbs, event types and inflections. In C.E. Johnson & J.H.V. Gilbert (eds.), *Children's language*, vol.ix. Mahwah, NJ: Erlbaum.

Clark, R. (1982). Theory and method in child language research: are we assuming too much?. In S.A. Kuczaj (ed.), *Language Development,* vol i: *syntax and semantics.* Hillsdale, NJ: Erlbaum.

Comrie, B. (1976). *Aspect.* Cambridge: Cambridge University Press.

Cziko, G. (1986). Testing the language bioprogram hypothesis: a review of children's acquisition of articles. *Language, 62,* 878-898.

De Houwer, A. (1990). *The acquisition of two languages from birth: a case study.* Cambridge: Cambridge University Press.

____ (1994). The separate development hypothesis: method and implications. In G. Extra & L. Verhoeven (eds.), *The crosslinguistic study of bilingual development,* pp. 39-50. Amsterdam, The Netherlands: Royal Netherlands Academy for Arts and Sciences.

____ (1995). Bilingual language acquisition. In P. Fletcher & B. MacWhinney (eds.), *The handbook of child language* , pp. 219-250. Oxford: Basil Blackwell.

____ (1997). The role of input in the acquisition of past verb forms in English and Dutch: Evidence from a bilingual child. In. E. Clark (ed.), *The Proceedings of the Twenty-eighth Annual Child Language Research Forum,* pp. 153-162. Stanford: Centre for the Study of Language Information.

Demuth, K. (1992). Accessing functional categories in Sesotho: interactions at the morhposyntactic interface. In J. Meisel (ed.), *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition,* pp. 83-108. Dordrecht: Kluwer.

Deprez, V. Pierce, A. (1993). Negation and functional projections in early grammars. *Linguistic Inquiry,* 23, 25-67.

Deuchar, M. (1997). Bilingual acquisition: Criteria for establishing one vs. two systems in early syntactic development. Paper presented at the I International Symposium on Bilingualism, University of Newcastle.

Deuchar, M. & Quay, S. (1998). One vs. two systems in early bilingual syntax: two versions of the question. *Bilingualism and Cognition, 1,* 231-243.

De Villiers, P.A. & De Villiers, J.G. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research, 2,* 331-341.

Donaldson, M. (1978). *Children's minds.* London: Fontana

Döpke, S. (1992). *One Parent one language: an interactional approach.* Amsterdam: John Benjamins.

____ (1997a). German-English simultaneous bilingualism: window into input processing during primary language acquisition. Paper presented at the International Symposium on Bilingualism, Newcastle.

____ (1997b). Is the simultaneous acquisition of two languages in early childhood equal to acquiring each of the two languages individually?. In E. Clark (ed.), *The Proceedings of the Twenty-eighth Annual Child Language Research Forum,* pp. 95-112. Stanford: Centre for the Study of Language Information.

____ (1998). Competing language structures: the acquisition of verb placement by bilingual German-English children. *Journal of Child Language, 25,* 555-584.

Emslie, H. & Stevenson, R. (1981). Preschool children's use of the articles in definite and indefinite referring expressions. *Journal of Child Language, 8,* 313-328.

Ervin-Tripp, S. (1964). Imitation and structural change in children's language. In E.H. Lenneberg (ed.), *New directions in the study of language,* pp. 163-189. Cambridge, Mass.: The Mit Press.

Fantini, A.E. (1978). Bilingual behaviour and social cues: case studies of two bilingual children. In M. Paradis (ed.), *Aspects of bilingualsim.* Columbia: Hornbeam Press.

Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S. & Reilly, J. (1994). *MacArthur Communicative Development Inventories: technical manual.* San Diego: San Diego State University.

Ferdinand, A. (1996). The development of functional categories. Dordrecht: ICG Printing.

Fernández-Martinez, A. (1994). El aprentizaje de los morfemas verbales: datos de uno estudio longitudinal. In S. López-Ornat, A. Fernández, P. Gallo & S. Mariscal (eds.), *La adquisición de la lengua española,* pp. 29-46. Madrid: Siglo XXI de España Editores.

Fillmore, C.J., Kay, P. & O'Connor, C. (1988). Regularity and idiomaticity in grammatical constructions: the case of Let Alone. *Language, 64,* 501-538.

Finch, S.P. & Chater, N. (1992). Bootstrapping syntactic categories. *Proceedings of the 14th Annual Conference of the Cognitive Science Society of America,* pp. 820-825. Bloomington, IN: Cognitive Science Society.

Fukui, N. (1986). *A theory of category projection and its applications.* Unpublished doctoral dissertation, MIT.

Garton, A. (1983). An approach to the study of determiners in early language development. *Journal of Psycholinguistic Research, 12,* 513-525.

Gathercole, V., Sebastian, E. & Soto, P. (1999). The early acquisition of Spanish verbal morphology: across the board or piecemeal knowledge? *International Journal of Bilingualism, 3.*

____ (in press). Lexically specified patterns in early verbal morhpology in Spanish. In M.R. Perkins & S.J. Howard (eds.), *New directions in language development and disorder.* London: Plenum.

Gawlitzek-Maiwald, I. & Tracy, R. (1996). Bilingual bootstrapping. *Linguistics, 34,* 901-926.

Genesee, F. (1989). Early bilingual development: one language or two?. *Journal of Child Language 16,* 161-179.

Genesee, F. , Nicoaldis E. & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language, 22,* 611-631.

Genesee, F., Boivin, I. & Nicoladis, E. (1996). Talking with strangers: A study of bilingual children's communicative competence. *Applied Psycholinguistics, 17,* 427-442.

Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure.* Chicago: University of Chicago Press.

Goodman, J. (1995). The shape of change: longitudinal evidence about language development. *Society for Research in Child Development Abstracts, 10,* 111.

Goodz, N. (1994). Interactions between parents and children in bilingual families. In F. Genesee (ed.), *Educating second language children,* pp. 61-81. New York: Cambridge University Press.

Gouze, K.R. & Nadelman, L. (1980). Constancy of gender identity for self and others in children between the ages of three and seven. *Child Development, 51,* 275-278.

Grimshaw, J. (1994). Minimal projection and clause structure. In Lust, B., Suñer, M. & Whitman, J. (eds.), *Syntactic theory and first language acquisition: crosslinguistic perspectives,* vol. i. *Heads, projections and learnability.* Hillsdale, NJ: Erlbaum.

Grinstead, J. (in press). Case, inflection and subject licensing in child Catalan and Spanish. *Journal of Child Language.*

Groat, E. (1995). English expletives: a minimalist approach. *Linguistic Inquiry, 26,* 354-365.

Guasti, M.T. (1993/94). Verb syntax in Italian child grammar: finite and non-finite verbs. *Language Acquisition, 3*: 1-40.

Guasti, M.T. (1996). Acquisition of Italian interrogatives. In H. Clahsen (ed.), *Generative Perspectives on Language Acquisition,* pp. 241-270. Amsterdam: John Benjamins.

Guasti, M.T. & Rizzi, L. (1996). Null Aux and the acquisition of residual V2. In *Proceedings of BUCLD 20,* 284-295.

Guilfoyle, E. & Noonan, M. (1988). Functional categories and language acquisition. Unpublished manuscript, McGill University, Montreal.

de Haan, G. (1986). De rol van morpfologie en syntaxis in de ontwikkeling van het werkwoord. *GLOT, 9,* 28-41.

Haegeman, L. (1994). Root infinitives, tense and truncated structures. *Language Acquisition,* 4.

Haider, H. & Prinzhorn, M. (eds.) (1986). *Verb second phenomena in Germanic languages.* Dordrecht: Foris.

Hamann, C. (1996). Null arguments in German child language. *Language Acquisition, 5:* 155-208.

Hamann, C. & Plunkett, K. (1996). Subject omission in child Danish. *Proceedings of BUCLD 20,* 220-232.

Hickey, T. (1993). Identfying formulas in first language acquisition. *Journal of Child Language, 20,* 27-41.

Hickmann, M., Hendriks, H., Roland, F.& Liang, J. (1996). The marking of new information in children's narratives: a comparison of English, French, German and Mandarin Chinese. *Journal of Child Language, 23,* 591-619.

Hoekstra, T. & Hyams, N. (1995). Missing heads in child language. In C. Koster & F. Wijnen (eds.), *Proceedings of GALA 1995.* Groningen: Centre for Language and Cognition.

Hoekstra, T. & Hyams, N. (1998). Aspects of root infinitives. *Lingua, 106,* 81-112.

Hoekstra, T., Hyams, N. & Becker, M. (1996). The underspecification of Number and the licensing of root infinitives. In A.Stringfellow, D. Cahan-Amitay, E. Hughes & A. Zukowski (eds.), *Proceedings of BUCLD 20,* 293-306.

Huang, J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry,* 15, 531-74.

Hulk, A. & van der Linden, E. (1997). Language mixing in a French-Dutch bilingual child. Paper presented at the I International Symposium on Bilingualism, University of Newcastle.

____ (1998). Evidence for transfer in bilingual children? *Bilingualism: Language and Cognition, 1,* 177-180.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M. & Lyons, T. (1991). Early vocabulary growth: relation to language input and gender. *Developmental Psychology, 27,* 236-248.

Hyams, N. (1986). *Language acquisition and the theory of parameters.* Dordrecht: Reidel.

____ (1987). The theory of parameters and syntactic development. In T. Roeper & E. Williams (eds.), *Parameter-setting and language acquisition.* Dordrecht: Foris.

____ (1992a). The genesis of clausal structure. In J. Meisel (ed.), *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition,* pp. 371-400. Dordrecht: Kluwer.

____ (1992b). Morphosyntactic development in Italian and its relevance to parameter-setting models: comments on the paper by Pizzuto & Caselli. *Journal of Child Language, 19,* 695-709.

____ (1994). V2, null arguments and COMP projections. In T. Hoekstra & B. Schwartz (eds.), *Language acquisition studies in generative grammar.* Amsterdam: John Benjamins.

____ (1996). The underspecifications of functional categories in early grammar. In H. Clahsen (ed.), *Generative perspectives on language acquisition,* pp 91-128. Amsterdam: John Benjamins.

Ingham, R. (1992). The optional subject phenomenonin young children's English: a case study. *Journal of Child Language, 19,* 133-151.

____ (1998). Tense without agreement in early clause structure. *Language Acquisition, 7,* 51-81.

Ingram, D. & Thompson, W. (1996). Early syntactic acquisition in German: evidence for the modal hypothesis. *Language,* 72, 97-120.

Jaeggli, O. & Hyams, N. (1988). Morphological Uniformity and the setting of the null subject parameter. *Proceedings of NELS,* 18, 238-253.

Jahn-Samilo, J. (1995). Langauge comprehension and production in children from 8 to 30 months of age: a comparison of parent report and laboratory measures from a longitudinal study. *Society for Research in Child Development, 10,* 112.

Jonas, D. (1994). The TP parameter in Scandinavian syntax. In *Gothenburgh Papers in Theoretical Linguistics, 70: Proceedings of the XIV Scandinavian Conference of Linguistics.* Gothenburgh: University of Gothenburg.

Jones, M. & Conti-Ramsden, G. (1997). A comparison of verb use in children with SLI and their younger siblings. *First Language, 17,* 165-193.

Jordens, P. (1990). The acquisition of verb placement in Dutch and German. *Linguistics, 28,* 1407-1448.

Jusczyk, P. (1981). Infant speech perception: a critical appraisal. In P. D. Eimas and J.L. Miller (eds.), *Perspectives on the study of speech.* Hillsdale N J: Erlbaum.

Jusczyk, P., Cutler, A. & Redanz, N.J. (1993). Preference for the predominant stress patterns of English words. *Child Development, 64,* 675-687.

Jusczyk, P., Friederici, A.D., Wessels, J.M.I., Svenkerud, V.Y. & Jusczyk, A. (1993). Infants' sensitivity to the patterns of native language words. *Journal of Memory and Language, 32,* 402-420.

Kail, M. & Hickmann, M. (1992). French children's ability to introduce referents in narratives as a function of mutual knowledge. *First Language, 12,* 73-94.

Karmiloff-Smith, A. (1979). *A functional approach to child language: a study of determiners and reference.* Cambridge: Cambridge University Press.

Katis, D. (1984). *The acquisition of the modern Greek verb with special reference to the imperfective past and perfect classes.* Unpublished doctoral dissertation, University of Reading.

Klima, E.S. & Bellugi, U. (1966). Syntactic regularities in the speech of children. In J. Lyons & R. Wales (eds.), *Psycholinguistics Papers,* pp. 183-207. Edinburgh: Edinburgh University Press.

____ (1985). Language and cognitive processes froma a developmental perspective. *Language and Cognitive Processes, 1,* 61-85.

Koehn, C. (1994). The acquisition of gender and number morphology within NP. In J. Meisel (ed.), *Bilingual first language acquisition: French and German grammatical development,* pp. 29-52. Amsterdam: John Benjamins.

Köpcke, K. (1988). Schemas in German plural formation. *Lingua, 74,* 303-335.

Köppe, R. (1996). Language differentiation in bilingual children: the development of grammatical and pragmatic competence. *Linguistics, 34,* 927-954.

Köppe, R. & Meisel, J. (1992). Code switching in bilingual first language acquisition. Unpublished manuscript, University of Hamburg.

Kramer, I. (1993). The licensing of null subjects in early child language. In *MIT Working Papers in Linguistics, 19,* 197-212.

Lanvers, U. (1997). New words before equivalents or viceversa? Vocabulary acquisition patterns in a bilingual child. Paper presented at the I International Symposium on Bilingualism, University of Newcastle.

Lanza, E. (1992). Can bilingual two-year-olds code switch? *Journal of Child Language, 19,* 633-658.

____ (1997). *Language mixing in infant bilingualism.* Oxford: Clarendon Press.

Lebeaux, D. (1987). Comments on Hyams. In T. Roeper & E. Williams (eds.), *Parameter-setting and language acquisition,* pp. 23-39. Dordrecht: Foris.

Leonard, L., Sabbadini, L., Volterra, V. & Leonard, J. (1988). Some influences on the grammar of English- and Italian-speaking children with specific language impairment. *Applied Psycholinguistics, 9, 39-57.*

Leonard, L.B., Bortolini, U., Caselli, M.C. & Sabbadini, L. (1993). The use of articles by Italian-speaking children with specific language impairment. *Clinical Linguistics and Phonetics, 7,* 19-27.

Leopold, W. (1949). *Speech development in a bilingual child,* vol. iii. New York: AMS Press.

Levy, Y (1975). Theoretical gains from the study of bilingualism: A case report. *Language Learning, 35,* 541-554.

____ (1983). It's frogs all the way down. *Cognition, 15,* 75-93.

____ (1988). On the early learning of formal grammatical systems: evidence from studies of the acquisition of gender and countability. *Journal of Child Language, 15,* 179-187.

Lieven, E. Pine, J. & Baldwin. G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language, 24,* 187-219.

Lindholm, K & Padilla, A. (1978). Child bilingualism: report on language mixing, switching and translations. *Linguistics, 211,* 23-44.

López-Ornat, S. (1997). What lies in between a pre-grammatical and a grammatical representation? Evidence on nominal and verbal form-function mappings in Spanish from 1;7 to 2;1. In A.T. Peerez-Leroux & W.R. Glass (eds.), *Contemporary perspectives on the acquisition of Spanish,* vol i. Cascadilla Press.

Lyon, I. (1997). Spontaneous production of past participles in Italian and underspecification of number. In Sorace, A., Heycock, C. & Shillcock, R. (eds.), *Proceedings of the GALA'97 Conference on Language Acquisition,* pp.96-101. Edinburgh: The University of Edinburgh.

Lyons, J. (1977). *Semantics,* vol i. Cambridge: Cambridge University Press.

MacWhinney, B. (1995). *The CHILDES Project: tools for analysing talk.* Second edition. Hillsdale, NJ: Erlbaum.

MacWhinney, B., Leinbach, J., Taraban, R. & McDonald, J.L. (1989). Language learning: cues or rules? *Journal of Memory and Language, 28,* 255-277.

MacWhinney, B., Bates, E. & Kliegel. R. (1984). Cue validity in sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behaviour, 23,* 127-150.

MacWhinney, B. & Snow, C.E. (1985). The child language data exchange system. *Journal of Child Language, 12,* 271-296.

Maratsos, M. (1976). *The use of definite and indefinite reference in young children.* London: Cambridge University Press.

Marchman, V. & Bates, E. (1994). Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language, 21,* 339-366.

Marcus, G.F., Ullman, M., Pinker, S., Hollander, M., Rosen, T.J. & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57* (Serial No 228).

Mayer, M. (1980). *Frog, where are you?*. Dial Books.

McCartney, K. (1984). Effect of quality of day care environment on children's language development. *Developmental Psychology, 2,* 244-260.

McClure, E. (1981). Formal and functional aspects of the code-switching discourse of bilingual children. In R. Durán (ed.), *Latino language and communicative behaviour*. Norwood, NJ: Ablex.

McDonald, L. & Pien, D. (1982). Mother conversational behaviour as a function of interactional intent. *Journal of Child Language, 9,* 337-358.

Mehler, J., Lambertz, G., Jusczyk, P. and Amiel-Tison, C. (1986). Discrimination de la langue maternelle pour le nouveau-né *Académie des Sciences 3,* 637-640.

Meisel, J. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam and Obler, L. (eds.), *Bilingualism across the lifespan: aspects of acquisition, maturity and loss.* Cambridge, Mass.: C.U.P.

Meisel, J. (1994). Getting FAT: Finiteness, Agreement and Tense in Early Grammars. In J. Meisel (ed.), *Bilingual first language acquisition: French and German grammatical development,* pp. 89-129. Amsterdam: John Benjamins.

Mervis, C.B. & Johnson, K.F. (1991). The acquisition of the plural morpheme: a case study. *Developmental Psychology, 27,* 222-235.

Mervis, C.B. & Bertrand, J. (1993). Early language and cognitive development: implications of research with children who have Williams Syndrome or Down Syndrome. Paper presented at the 60th Anniversary Meeting of the Society for Research in Child Development. New Orleans, LA.

Mills, A. (1986). *The acquisition of gender: a study of English and German.* Berlin: Springer-Verlag.

Müller, N. (1994). Gender and number agreement within DP. In J. Meisel (ed.), *Bilingual frst language acquisition: French and German grammatical development,* pp. 53-89. Amsterdam: John Benjamins.

____ (1998). Transfer in bilingual first language acquisition. *Bilingualism: Language and Cognition, 1,* 151-192.

Nicoladis, E. (1994). *Code-mixing in young bilingual children.* Unpublished doctoral dissertation, Mc Gill University.

Ochs, E. (1979). Transcription as theory. In Ochs & B Schieffelin (eds.), *Developmental pragmatics.* New York: Academic Press.

Olguin, R. & Tomasello, M. (1993). Twenty-five-month-old children do not have a grammatical category of verb. *Cognitive Development, 8,* 245-272.

Orsolini, M., Rossi, F. & Pontecorvo, C. (1996). Re-introduction of referents in Italian children's narratives. *Journal of Child Language, 23,* 465-486.

Paradis, J. & Genesee, F. (1996a). On continuity and the emergence of functional categories in bilingual first language acquisition. Manuscript, McGill University, Montreal.

____ (1996b). Syntactic acquisition in bilingual children: autonomous or interdependent?. *Studies in Second Language Acquisition, 18,* 1-25.

Pearson, B. , Fernández, S.C. & Oller, D.K. (1995). Cross-language synonyms in the lexicons of bilingual infants: one language or two? *Journal of Child Language, 22,* 345-368.

Pearson, B., Fernández, S.C., Lewedeg, V. & Oller, D.K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18,* 41-58.

Pellegrino, M.L. & Scopesi, A. (1990). Structure and function of baby talk in a day-care centre. *Journal of Child Language, 17,* 101-114.

Pérez-Pereira, M. (1991). The acquisition of gender: what Spanish children tell us. *Journal of Child Language, 18,* 571-590.

Peters, A.M. (1983). *The units of language acquisition.* Cambridge: Cambridge University Press.

Piaget, J. (1926). *The language and thought of the child.* New York: Harcourt.

____ (1968). *On the development of memory and identity.* Worcester, Mass.: Lark University Press.

Pierce, A. (1992). *Language acquisition and syntactic theory: a comparative analysis of French and English child grammars.* Dordrecht: Kluwer.

Pine, J. (1990). Individual differences in early language development and their relationship to maternal style. Unpublished doctoral dissertation, University of Manchester.

Pine, J. & Martindale, H. (1996). Syntactic categories in the speech of young children: the case of the determiner. *Journal of Child Language,23,* 369-395.

Pine, J. & Lieven, E. (1993). Reanalysing rote-learned phrases; Individual differences in the transition to multi-word speech. *Journal of Child Language, 20,* 551-571.

____ (1997). Slot-and-frame patterns in the development of tbe determiner category. *Applied Psycholinguistics, 18,* 123-138.

Pine, J., Lieven, E. & Rowland, C. (1998). Comparing different models of the development of the English verb category. *Linguistics, 36:* 807-830.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, Mass.: Harvard University Press.

____ (1991). Rules of language. *Science, 253*, 530-535.

Pizzuto, E., Caselli, M.C. & Tangorra, T. (1991). The acquisition of Italian definite articles: insights from the study of natural language and neural networks. Paper presented at the 16th BUCLD.

Pizzuto, E. & Caselli, M.C. (1992). The acquisition of Italian morphology: implications for models of language development. *Journal of Child Language, 19*, 491-557.

____ (1994). The acquisition of Italian verb morphology in a cross-linguistic perspective. In Y. Levy (ed.), *Other children, other languages,* pp. 137-187. Hillsdale, NJ: Erlbaum.

Platzack, C. (1990). A grammar without functional categories: a syntactic study of early Swedish child language. *Nordic Journal of Linguistics, 13,* 107-126.

____ (1992). Functional categories and early Swedish. In J. Meisel (ed.), *The acquisition of verb placements: functional categories and V2 phenomena in language development.* Dordrecht: Kluwer.

Plunkett, K. (1993). Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language, 20,* 43-60.

Poeppel, D. & Wexler, K. (1993). The Full Competence Hypothesis of clause structure in Early German. *Language, 69*, 1-33.

Pollard, C. & Sag, I. (1994). *Head-driven phrase structure grammar.* Chicago: University of Chicago Press.

Pollock , J-Y (1989). Verb movement, UG and the structure of IP. *Linguistic Inquiry, 20,* 365-424.

Power, R.J.D. & Dal Martello, M.F. (1986). The use of definite and indefinite articles by Italian preschool children. *Journal of Child Language, 13,* 145-154.

Quay, S. (1995). The bilingual lexicon: implications for the study of language choice. *Journal of Child Language, 22,* 369-387.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax.* Oxford: Blackwell.

____ (1992). The acquisition of the morphosyntax of finite verbs in English. In J. Meisel (ed.), *The acquisition of verb placements: functional categories and V2 phenomena in language development,* pp. 23-62. Dordrecht: Kluwer.

____ (1994). The syntax of questions in child English. *Journal of Child Language, 21*, 211-236.

____ (1995). Phrase structure and functional categories. In P. Fletcher & B. MacWhinney (eds.), *The handbook of child language,* pp. 483-507. Oxford: Blackwell.

____ (1996). Towards a structure-building model of acquisition. In H. Clahsen (ed.), *Generative perspectives on language acquisition,* pp. 43-90. Amsterdam: John Benjamins.

Raduloviç, L. (1975). *Acquisition of language: studies of Dubrovnik children.* Unpublished doctoral dissertation, University of California, Berkeley.

Redington, M. & Chater, N. (1998). Connectionist and statistical approaches to language acquisition: a distributional perspective. 129-191.

Redlinger W. E. and Park, T. (1980). Language mixing in young bilinguals. *Journal of Child Language, 7,* 337-352.

Rispoli, M. & Bloom, L. (1985). Incomplete and continuing: theoretical issues in the acquisition of tense and aspect. *Journal of Child Language, 12,* 471-474.

Rizzi, L. (1982). *Issues in Italian Syntax.* Dordrecht: Foris.

____ (1986). Null objects in Italian and the theory of *pro. Linguistic Inquiry, 17,* 501-557.

____ (1991). Residual verb second and the wh-criterion. Manuscript, University of Geneva.

____ (1994a). Early null subjects and root null subjects. In T. Hoekstra & B. Schwartz (eds.), *Language acquisition studies in generative grammar.* Amsterdam: John Benjamins.

____ (1994b). Some notes on linguistic theory and language development: the case of root infinitives. *Language Acquisition , 3,* 371-93.

____ (1997). Residual verb second and the wh-criterion. In A. Belletti & 1. Rizzi (eds.), *Parameters and functional heads: essays in comparative syntax,* pp. 63-90. Oxford: Oxford University Press.

Roeper, T. (1996). The role of merger theory and formal features in acquisition. In H. Clahsen (ed.), *Generative perspectives on language acquisition,* pp 129-160. Amsterdam: John Benjamins.

Rollins, P. & Snow, C.E. (1998). Shared attention and grammatical development in typical children and children with autism. *Journal of Child Language, 25,* 653-673.

Ronjat, J.A. (1913). *Le development du language observé chez un enfant bilingue.* Paris: Champion.

Rubino, R. & Pine, J. (1998). Subject-verb agreement in Brazilian Portuguese: what low error rates hide. *Journal of Child Language, 25,* 35-59.

Safir, K. (1993). Perception, selection and structural economy. *Natural Language Semantics, 2,* 47-70.

Sano, T. (1995). *Roots in language acquisition: a comparative study of Jspanese and European Languages.* Unpublished doctoral dissertation, Uiversity of California, Los Angeles.

Sano, T. & Hyams, N. (1994). Agreement, finiteness and the development of null arguments. *Proceedings of NELS, 24,* 543-558.

Schaeffer, J. (1997). *Direct object scrambling in Dutch and Italian child language.* Unpublished doctoral dissertation, University of California, Los Angeles.

Schaffer, H.R. & Liddell, C. (1984). Adult-child interaction under dyadic and polyadic conditions. *British Journal of Developmental Psychology, 2,* 33-42.

Schaffer, H.R. & Crook, C.K. (1979). Maternal control techniques in a directed play situation. *Child Development, 50,* 989-998.

Schelletter, C. & Sinka, I. (1997). Latvian-English and German-English bilingual acquisition: new light on Universal Grammar. Paper presented at the First International Symposium on Bilingualism: Bilingual communities and individuals, Vigo.

Schütze, C. & Wexler, K. (1996). Subject case licensing and English root infinitives. *Proceedings of BUCLD, 20,* 670-81.

Shirai, Y. & Andersen, R.W. (1995). The acquisition of tense-aspect morphology: a prototype account. *Language, 71,* 743-762.

Singer Harris, N., Bellugi, U., Bates, E., Rossen, M. & Jones, W. (1997). Emerging language in two genetically based neurodevelopmental disorders. In D. Thal & J. Reilly (eds.), Special issue on Origins of Communication Disorders. *Developmental Neuropsychology, 13,* 345-370.

Sinka, I. & Schelletter, C. (1998). Morphosyntactic development in bilingual children. *International Journal of Bilingualism, 2,* 301-326.

Smith, C. (1983). A theory of aspectual choice. *Language, 59,* 479-501.

____ (1991). *The parameter of aspect.* Dordrecht: Kluwer.

Snow, C.E. (1989). Understanding social interactions and language acquisition: sentences are not enough. In M. Bornstein & J. Bruner (eds.), *Interactions in human development,* pp. 83-103. Hillsdale, NJ: Erlbaum.

Snyder, L., Bates, E. & Bretherton, I. (1981). Content and context in early lexical development. *Journal of Child Language, 8,* 565-582.

Stromswold, K. (1990). *Learnability and the acquisition of auxiliaries.* Unpublshed doctoral dissertation, MIT.

____ (1995). The acquisition of subject and object wh-questions. *Language acquisition, 4,* 5-48.

Swain, M. (1977). Bilingualism, monolingualism and code acquisition. In W. Mackey and T. Andersson (eds.), *Bilingualism in early childhood.* Rowley MA: Newbury House.

Swain, M. & Wesche, M. (1975). Linguistic interaction: case study of a bilingual child. *Language Sciences, 37,* 17-22.

Sylva, K., Roy, C. & Painter, M. (1980). *Child watching at playgroup and nursery school.* London: Grant MacIntyre.

Taeschner, T. (1983). *The sun is feminine: a study in language acquisition in bilingual children.* Berlin: Springer-Verlag.

Thal, D., Bates, E. & Bellugi, U. (1989). Language and cognition in two children with Williams syndrome. *Journal of Speech and Hearing Research, 3,* 489-500.

Thal, D., Bates, E., Goodman, J. & Jahn-Samilo, J. (1997). Continuity of language abilities in late- and early-talking toddlers. In D. Thal & J. Reilly (eds.), Special issue on Origins of Communication Disorders. *Developmental Neuropsychology, 13,* 239-273.

Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics, 10,* 335-355.

Tizard, B., Carmichael, H., Hughes, M. & Pinkerton, G. (1980). Four-year-olds talking to mothers and teachers. In L.A. Hersov & M. Berger (eds.), *Language and language disorders in childhood.* Oxford: Pergamon Press.

Tizard, B., Phillips, J. & Plewis, J. (1976). Staff behaviour in preschool centres. *Journal of Child Psychology and Psychiatry, 17:* 21-33.

Tomasello, M. (1992). *First verbs: a case study of early grammatical development.* Cambridge: Cambridge University Press.

Tomasello, M. & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language, 4,* 197-212.

Tomasello, M. & Kruger, A.C. (1992). Joint attention on actions: acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19,* 311-333.

Tomasello, M. & Olguin, R. (1993). Twenty-three-month-old children have a grammatical category of noun. *Cognitive Development, 8,* 451-464.

Tomasello, M. Strosberg, R. & Akhtar, N. (1996). 18-month old children learn words in non-ostensive contexts. *Journal of Child Language, 23,* 157-176.

Trehub, S. (1973). Auditory-linguistic sensitivity in infants. Unpublished doctoral dissertation, Mc Gill University, Montreal.

Tsimpli, I. M. (1996). *The prefunctional stage of language acquisition: a crosslinguistic study.* New York: Garland.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition, 40,* 21-81.

____ (1992). Categories of first syntax: be, be+ing, and nothingness. In J. Meisel (ed.) *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition* , pp. 401-422. The Netherlands: Kluwer.

Valian, V. & Eisenberg, Z. (1996). The development of syntactic subjects in Portuguese-speaking children. *Journal of Child Language, 23,* 103-128.

Van Gelderen (1993). *The rise of functional categories.* Amsterdam: John Benjamins.

Vendler, Z. (1967). *Linguistics in philosophy.* Ithaca, NY: Cornell University Press.

Verrips, M. & Weissenborn, J. (1992). Routes to verb placement in early German and French: the independence of finitiness and agreement. In J. Meisel (ed.), *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition,* pp. 371-400. Dordrecht: Kluwer.

Vihman, M. (1985). Language differentiation by the bilingual child. *Journal of Child Language, 12,* 297-324.

Vihman, M. & Mc Cune, L. (1994). When is a word a word? *Journal of Child Language, 21,* 517-542.

Volterra, V. (1976). A few remarks on the use of the past participle in child language. *Italian Linguistics, 2,* 149-157.

Volterra, V. and Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Langauge, 5,* 311-326.

Warden, D. (1976). The influence of context on children's use of identifying expressions and references. *British Journal of Psychology, 67,* 101-112.

Watanabe, A. (1996). *Case checking and wh-agreement.* Dordrecht: Kluwer.

Weissenborn, J. (1992). Null subjects in early grammars: implications for parameter-setting theory. In J. Weissenborn, H. Goodluck & T. Roeper (eds.), *Theoretical issues in language acquisition: continuity and change in development,* pp. 269-99. Hillsdale, NJ: Lawrence Erlbaum Associates.

Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. In D. Lightofoot & N. Hornstein (eds.), *Verb Movement,* 305-350. Cambridge: Cambridge University Press.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua, 106,* 23-79.

Wigglesworth, G. (1990). Children's narrative acquisition: a study of some aspects of reference and anaphora. *First Language, 10,* 105-125.

Appendix 1

Questionnaire

In the attempt to get a clearer picture of the kind of linguistic input your child receives I am asking you to fill this questionnaire. Some background information on your child's first two years of life will also be required to assess the more general pattern of exposure to English and Italian.

Please note that the questionnaire is simply intended as a prompt, feel free to add more information or to make any changes you think are appropriate.

Your child's (C.) first two years

1. C.'s date of birth:

2. C.'s place of birth:

3. C.'s family, their native languages and languages used with C.:

4. What other people did C. come into regular contact with in his/her first two years of life? State their native language and languages used with C.:

5. Was C. exposed to both languages within the first week of life?

6. Has exposure to the two languages changed significantly in C.'s first two years? If yes state when and how it changed:

7. What child care provisions have been made for C. in the last two years?

8. When did C. start to attend nursery?

9. Has the number of hours C. spends at the nursery changed over the last two years? If yes state how and when:

10. How many trips to Italy has C. taken over the last two years? How long was he there for each time?

11. With the help of the checklist below state all of C.'s sources of input for both languages (mark **E** for English and **I** for Italian where appropriate):

adult monolingual speakers

child monolingual speakers

adult bilingual speaker

child bilingual speakers

books (picture books)

audiovisual material (videos, films, TV)

songs/nursery rhymes

12. What is the parents' strategy in bilingual contexts?

a. switch between Italian to C. and English to the other speakers
b. switch to English to C. in order to include the English speakers in the conversation

A week-day in the life of your child

13. morning routine

Briefly describe C.'s daily getting-up routine including information on :

- parent(s) who wake(s) C. up and get(s) him/her ready for breakfast
- people present at breakfast and languages spoken
- whether TV, radio or video are on in the morning before going out for the day
- parent(s) who take(s) C. to the nursery

14. at the nursery (to be completed by caretakers at the nursery)

Briefly describe the group C. is part of including information on:

- children's age and linguistic development (e.g. one/two-word stage)
- number of caretakers per group
- activities children engage in (e.g. singing, playing games, looking at books)
- routines such as nappy-changing, eating , taking a nap, getting ready to leave, etc.

15. lunch time

Briefly describe a typical lunch break during the week including information on:

- parent/caretaker collecting C. from nursery
- people present at lunch time and languages spoken
- whether TV, radio or video are on

16. in the afternoon
Give information on C.'s afternoons including details on:

- people looking after him and languages spoken
- other people present (e.g. siblings, other children or adults)
- activities (e.g. games, TV, videos, music, going out, etc.)

17. In the evening
Give details on:

- time when parent(s) come(s) home
- activities with parents/siblings before the evening meal
- visitors and languages spoken
- whether TV, video or radio are on during the evening

18. bed-time routine
Briefly describe C.'s bed-time routine including information on:

- bath-time:
- whether stories are read to C. before going to bed: no bed-time stories.

Week-ends

Family life at the week-end will presumably be different from the weekly routine, Try to give exhaustive information on:

- regular week-end activities (e.g. church attendance, visits to relatives/friends, playgroups, etc.)
- shopping
- cinema, theatre (e.g. pantomimes, puppets), children's parties
- trips away from home

19. Give an approximate percentage estimate of C.'s exposure to the following linguistic contexts:

a. strictly monolingual Italian context
b. stricly English monolingual context

c. English- Italian bilingual context
d. monolingual Italian context with occasional English spoken
e. monolingual English context with occasional Italian spoken

20. What is the parents' attitude towards codeswitching?

a. parents ignore mixed utterances and continue the conversation in the language it was started in.
b. parents ask the child to repeat pretending not to understand the mixed utterance.
c. parents explicitly ask the child to use only one language at a time.
d. Are parents more likely to accept intrasentential code-switching where a lexical (or possibly a functional) item is "borrowed" from the other language (e.g. Ho visto un *dog*, I *vado* to school); or intersentential code-switching where the child switches back and forth from one language to other in a more extensive way?

21. What's the parents' own use of code-switching? Do they try to eliminate it from their speech addressed to the children?

Thank you for your cooperation!

Appendix 2

Transcription and coding systems

In this section detailed information is given on the CHAT transcription and coding of the videorecorded material in the Carlo corpus

The three major components of a CHAT transcript are the file headers, the main tier and the dependent tiers. File headers are lines of text preceded by the "@" sign and contain information such as name and age of participants, the date of the recording, the language(s) spoken, information about the situation's setting, etc. Headers are divided into obligatory, constant and changeable. There are only three obligatory headers in a CHAT transcript that must always appear: @Begin, @Participants and @End. The @Begin and the @End headers are there to limit the upper and lower boundaries of the transcript and to ensure that no material is accidentally lost or added. The @Participants header gives the names and the roles of the various people involved in the recording and is essential to subsequently identify all the participants within the file. The information contained in the constant headers applies throughout the file, while the information in the changeable headers may apply only to some portions of the transcript. A changeable header such as @Activites can be useful to describe in more detail the specific actions performed by the participants at some particular points of the transcript.

In addition to the obligatory headers (@Begin; @Participants; @End), C.'s transcripts contain five constant headers placed at the beginning of the file: @Filename; @Date; @Age of CAR; @Situation; @Language as in example (12) below:

(1)

@Begin
@Filename: carloita1.cha
@Participants: CAR Carlo Target_Child, ELE Elena Family_Friend
@Date: 19-SEP-1996
@Age of CAR: 1;10
@Situation: looking at picture books
@Language: Italian

The speaker's utterances are transcribed on what is known in CHAT format as the main line. The main line begins with an asterisk and is followed by a three block letter speaker ID, a colon and a tab:

(2) File carloita1.cha

*ELE: e cos' è questa?
%eng: and what is this?
*CAR: ina.
%eng: small.
*ELE: bravo, è una tartarughina!
%eng: good it's a small turtle.

Each line contains one and only one utterance delimited by an utterance terminator: a period, a question mark or an exclamation mark. An utterance is here loosely defined as a stretch of speech with the same intonational contour. Utterances are made up of one or more words, and if deciding what constitutes a word in the adult language is relatively straightforward, the same cannot be said for child language where many of the children's early "words" are morphologically and phonologically very different from the adult target. In establishing what counts as a word I followed four of the five criteria devised by Deuchar and Quay (1998: 6):

- the reference or use had to be similar to the adult's, or to the child's own in other situations;
- there had to be phonetic similarity to a source word; for example, it might be a phonetically simple version of an adult word
- the adult's behaviour must not i ndicate that they have not recognized the word (except when the word comes from a language not currently being used by the adult
- the meaning of any accompanying non-verbal behaviour must not conflict with the meaning of the word" (Deuchar & Quay, 1998: 6)

A fifth criterion was to exclude imitated words, a decision dictated by the specific research question that Deuchar & Quay were adressing, i.e. the emergence of translation equivalents in the emergent lexicon of a bilingual child. For my own research purposes I did not intend to exclude imitated words and therefore followed Vihman & McCune (1994) in including imitations in order to help distinguish words from babble.

All the data was transcribed orthographically, although a phonetic transcription was occasionally provided for some child forms together with the corresponding adult target. Pauses, overlaps, retracings, tag questions, hesitations, self-repetitions, interruptions, self-interruptions, trailing offs, quotations, self-completions and other completions are all marked on the main line using a series of symbols provided by

CHAT[1]. Unintellegible words were marked as "xxx", while untranscribed material, such as asides or longer conversations where C. was not included, were simply marked as "www" with the addition of an %exp dependent tier with the reason why it was not transcribed:

(3) File carloita11.cha

*LUD: www.
%exp: makes arrangements with the childminder.

Second language forms, such as, for example, Italian words used in an English session, were coded using the special marker "@s" so that it would later be possible to assess C.'s language choice and his competence in addressing speakers in the appropriate language.

No morphemicization was performed on the main line; for the English data, a %mor dependent tier with a complete morphological tagging of all utterances was added after completing the data transcription.

In the spirit of the guidelines for transcribers set by MacWhinney (1995), I tried to produce a clear, systematic and readable transcription of the data collected in a format that would be computer readable and accessible to other researchers. It goes without saying that turning spoken data into written data inevitably imposes a selection process whereby the transcriber decides *what* to transcribe and *how* to transcribe it. Written transcripts are normally considered the "raw data" which the researcher can draw from to test the validity of his or her hypotheses. It is worth remembering that they are the mediated result of the transcriber's work and as such the outcome of a more or less deliberate manipulation. As Ochs (1979) notes in an influential article on transcription as theory, "transcription is a selective process reflecting theoretical goals and definitions" (Ochs, 1979: 44). She urges awareness of the filtering process to which the transcriber subjects the data, and she notes that "[s]electivity, then, is to be encouraged" (Ochs, 1979: 44). Different researchers will have different agendas, and they must keep these in mind when transcribing their data. In this case the initial focus of the research project was on the emergence of morphosyntax in C.'s two languages, therefore particular attention was paid to an accurate transcription of the child's

[1] See MacWhinney (1995) for a comprehensive description of transcription conventions on the main line.

utterances, including self-corrections, hesitations and pauses that may be somehow revealing of the child's competence (or lack thereof).

Transcribing child data is a time-consuming activity at the best of times, i.e. when one simply focuses on the transcription of actual speech. Providing additional information, and coding it accordingly on relevant dependent tiers clearly requires extra time and in every research project there are inevitably time restrictions one must be aware of. Dependent tiers are lines below the text on the main line that include codes, comments, descriptions of interst to the researcher. MacWhinney (1995) lists twenty-three standard dependent tiers with the proviso that others can be created by the transcriber to suit his or her own specific needs. All dependent tiers begin with the "%" symbol followed by a colon and a tab; both the tier code and the text should be in lower case. Except for the %mor and the %syn tiers these lines, unlike the main lines, do not require ending punctuation.

In C.'s transcripts in addition to transcribing the main line, adding information on pauses, overlaps, retracings, tag questions, hesitations, self-repetitions, interruptions, self-interruptions, trailing offs, quotations, self-completions and other completions, I also added a %mor tier for the entirety of the English data, and %err tier where relevant for specific phonological errors, and a %gpx tier to code gestural and proxemic material when I felt it would contribute significantly to the utterance it referred to.

Since the core of the research project was going to be on the emergence of morphosyntax, the generation of a %mor tier proved extremely useful for subsequent targeted morphological and syntactic analyses. Unfortunately, because of limitations to the CLAN program MOR, generation of the %mor tier was, at the time of writing, only available for English. As a consequence, only the English data in C.'s corpus are morphologically tagged by %mor. On the %mor line words are tagged by their part of speech or "scat", followed by the separator "|" which is in turn followed by the word itself broken down into the various morphemes:

(4) File carloeng16.mor

*ERI: oh you're running out of space Carlo .
%mor: co|oh pro|you-v:aux|be&PRES v|run-PROG prep|out prep|of n|space n:prop|Carlo .

Unlike the %mor tier, which is a complete tagging of the main line in the English data, I only used the %err and the%gpx tiers occasionally whenever some useful information needed to be coded independently of the main line.

I used the %err tier to categorise and analyse errors marked on the main line by the [*] symbol. On the main line the transcriber places the [*] symbol after the error, then on the %err tier the erroneous form is repeated, in the case of phonological errors it is followed by a phonemic transcription in UNIBET (an ASCII translation of IPA symbols used by CHILDES for single character phonemic transcription), in turn followed by the "=" symbol, by the target form and, in the case of phonological errors, by the phonemic transcription in UNIBET.

The two main types of errors I marked were phonological errors and language choice errors. Phonological errors were all those errors involving specific phonological units, where the child used a form that was recognisable in the adult target but not standard (cf. Deuchar & Quay's criteria for deciding what constitutes a word in child language). The error is also coded according to its source, i.e. it could be phonological, syntactic, intonational, morphological, etc. MacWhinney (1995) lists a number of general codes, type codes, phonological codes, morphological and syntactic codes that may be used to identify the nature of the error in question.

(5) File carloeng1.mor

*KAR: what's that?
%mor: wh:pro|what-v|be&3S pro:dem|that?
*CAR: raffe [*].
%mor: n|raffe.
%err: raffe /raf/ = giraffe /dZ6raf/ $PHO

In (5) above C. produces the truncated form /raf/ instead of the adult target /dz6raf/. The error is marked on the main line by the [*] symbol and a full analysis and categorisation is provided on the %err tier. Because the error involves a nonstandard phonological form the UNIBET transcription is provided both for the erroneous form and for the target form, and the error is appropriately identified by the $PHO symbol as a phonological error.

As for language choice errors, these indicate the pragmatically inappropriate use of language A when the required language, given the addressee's language choice, would be language B. In the example below, a word from Italian was used by C. with

an English monolingual speaker when the English equivalent would have been required instead:

(6) File carloeng1.mor

*KAR: you know what that is.
%mor: pro|you v|know wh:pro|what pro:dem|that v|be&3S
*CAR: palla@s [*].
%mor: n|palla
%err: palla = ball $LAN
*KAR: that's in Italian.
%mor: pro:dem|that-v|be&3S prep|in n|Italian .
*KAR: what is it in English ?
%mor: wh:pro|what v|be&3S pro|it prep|in n|English ?
*CAR: palla@s [*].
%mor: n|palla.
%err: palla = ball $LAN

The language choice here is clearly inappropriate; instead of providing the required English noun "ball" C. produces the Italian equivalent "palla". The English interlocutor signals this breakdown in communication by pointing out the wrong language choice and prompting C. to provide the English noun "ball", but without success. Like in the previous example the word is marked by an asterisk to indicate it is an error and the special marker @s indicating that it is a second language form different from the main language used throughout the rest of the transcript, on the %err tier the word is repeated followed by the "=" sign, by the correct English translation, and finally by the $LAN error code to indicated that it is a language choice error type. Note, however, that not all second language forms coded by the special marker @s must be treated as language choice errors. This is the case only if the child's use of the second language form is inappropriate in that context because the addressee, the topic or the setting would require the child to use another language instead.

(7) File carloeng18.mor

*CAR: <le@s mie@s> [<] sono@s calde@s e@s fredde@s [% talking to Ludovica] .
%eng: mine are hot and cold.
%mor: det|le pro:poss|mie v|sono adj|calde conj:coor|e adj|fredde .
*LUD: sono@s calde@s e@s fredde@s ?
%eng: are they hot and cold?
%mor: v|sono adj|calde conj:coor|e adj|fredde ?

Although the language used throughout most of the transcript from which the example above is taken is English, since it is a session where C.'s main interlocutor is E., an English speaker, C. appropriately uses Italian when addressing myself. Later on

the child appropriately switches to using English with E. while I was still in the room, thus showing that he is indeed capable of choosing the language appropriately according to the addressee:

(8) File carloeng18.mor

*CAR: I found a bit of the train .
%mor: pro|I v|find&PAST det|a n|bit prep|of det|the n|train .
*ERI: you have .
%mor: pro|you v:aux|have .

The third dependent tier appearing in C.'s corpus is the %gpx tier. It codes for gestural and proxemic material, such as nodding, reaching, pointing, etc. On certain occasions I found it useful to include such information, in order to complement the verbal context with nonverbal information. As Ochs (1979) emphasises repeatedly, children, like adults, rely on gestures, body orientation, and eye gaze to communicate effectively. In fact, the younger children are, the more likely they are to make use of nonverbal behaviour to get their message across. At times it may even be the case that nonverbal behaviour is not so much an accompaniment to verbal behaviour, as an alternative to it.

(9) File carloeng4.mor

*CAR: that ?
%mor: pro:dem|that ?
%gpx: showing a wooden block to Karen

Here C. uses a gesture to accompany his question, while in the example below he complies with E.'s request by pointing rather than by speaking:

(10) File carloeng14.mor

*ERI: where are the fish ?
%mor: wh:adv|where v|be&PRES det|the n|fish ?
*CAR: 0.
%mor: .
%gpx: points to the card he's holding .

In itself C.'s failure to reply verbally to E.'s question provides several pieces of information. On this occasion C. has decided to rely on gestural information, an option which is perfectly acceptable given the situational context: he is sitting facing his

interlocutor and the card with the fish on is clearly visible to both of them. Using a deictic gesture such as pointing is perfectly sufficient to signal that he has understood the question and that he can give the correct answer. Secondly, the fact that he is pointing, although he does not say anything, also tells us that he is willing to comply with the expectations of his interlocutor and is prepared to acknowledge his question. Not saying anything, marked on the main line as "0', and not doing anything would have meant ignoring the request altogether and would have led to a breakdown in the communication. Although C. does not say anything he does something: by resorting to nonverbal behaviour he manages to keep the exchange going. Identifying this kind of information by using the %gpx tier is very useful in the assessment of the child's communicative competence, and it is an additional source of information on the child's use of his two languages.

Appendix 3

Italian verbs

The following table lists all of C.'s Italian verbs by recording session. The verbs are listed in alphabetical order and for each of them there is a chronological record of how many times a particular verb form for a given verb type appears in a file. The last two rows show the number of verb types per session, the number of verb types used contrastively in each session, and the percentage of verb types used contrastively in each session.

		File numbers																		
V type	V forms	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
AIUTARE	2 p.s. Pres.																		1	1	
to help	1 p.s. Pres.																			1	
ANDARE	Inf.													3		9				6	1
to go	3 p.s. Pres.										1							1		15	3
	3 p.s. Pres. Perf.												1						1		1
	3 p.s. Pres. Prog.													1							
	1 p.s. Pres.														1						
	Bare P.P.														2				1		
	3 p.p. Pres.																			3	8
	2 p.s. Pres. Perf.																			2	
	3 p.s. Imperf.																				3
APRIRE	3 p.s. Pres.		1									1						1			
to open	2 p.s. Imp. Aff.							2													
	1 p.p. Imp. Aff.																				1
ARRIVARE	3 p.s. Pres.													2							
to arrive																					
AVERE	3 p.s. Pres.										2	2			3			2			
to have	1 p.s. Imperf.																1				
	3 p.p. Pres.																		2	1	1
	Inf.																			1	
	1 p.s. Pres.																				2
	3 p.s. Imperf.																				1
BALLARE	3 p.s. Pres.												1								
to dance	2 p.p. Pres. Perf.														2						
BATTERE	Inf.															1					
to hit	3 p.s. Pres. Perf.															1					
BERE	3 p.s. Pres. Perf.														1						
to drink	2 p.s. Pres.														1						
BUTTARE	1 p.s. Pres.												1	1							
to throw away	1 p.s. Pres. Perf.														1						
	3 p.s. Pres.																		1		
CADERE	Bare P.P.							2						1					1		
to fall	3 p.s. Pres. Perf.												1	4		2			2		1
	3 p.p. Pres.													1							
	3 p.p. Pres. Perf.													7		2					

CAMMINARE to walk	3 p.s. Pres. Prog. 3 p.s. Pres. Inf.													5	1 1	 1					
CASCARE to fall	3 p.s. Pes. Prog.															2					
CERCARE to look for	1 p.s. Pres.			1						2									1		
CHIAMARSI to be called	3 p.s. Pres. 1 p.s. Pres.																			3	 1
CHIUDERE to close	2 p.s. Imp. Aff. 1 p.s. Pres. 2 p.s. Pres. 1 p.p. Pres.		1				1	1							2		4 1				
COLORARE to colour	1 p.s. Pres.																		1		
CORRERE to run	3 p.s. Pres. 3 p.s. Pres. Prog. 3 p.s. Imperf.		1	1														1			2
DARE to give	23 p. p.s. Pres. 2 p.s. Pres. Perf. 3 p.s. Pres. Perf. 1 p.s. Pres. 3 p.s. Pres.														1	1		1	2 1 1		1
DIRE to say	3 p.s. Pres. Perf. 1 p.s. Pres. 1 p.s. Pres. Perf. 2 p.s. Pres. Perf. Inf. Bare P.P. 3 p.s. Pres. 2 p.s. Pres. 2 p.s. Imperf.																	3	1	1 2 2	1 1 1 1 2 1 1
DISEGNARE to draw	3 p.s. Pres. Prog.														1						
DIVERTIRSI to have fun	1 p.s. Pres. Perf.																	1			
DONDOLARE to swing	3 p.p. pres. Prog.													3							

DORMIRE to sleep	3 p.s. Pres.																		2		
DOVERE must	2 p.s. Pres. 3 p.s. Pres. 1 p.s. Pres. 1 p.p. Pres. 3 p.p. Pres.															1 3	 1	2	3 2	2 1 3	2 1 4 1
ENTRARE to enter	Inf. 3 p.s. Pres.															4 2					
FARE to do/to make	3 p.s. Pres. 1 p.s. Pres. 1 p.p. Pres. 3 p.p. Pres. 2 p.s. Pres. 3 p.s. Pres. Perf. Bare P.P. 1 p.s. Pres. Perf. Inf. 2 p.s. Imp. Aff. 1 p.p. Pres. Perf. 3 p.p. Pres. Perf. 3 p.s. Imperf.			1				1		1	4	1	16	23 1 1 1 1 2	4 1 1 10 2	2 1 6 1 1 2 1	 1 1 2 2	3 1 4 1 1	4 1 3 3 3 4 1	 3 1 3 	 2 2 2 1
FINIRE to finish	2 p.s. Pres. Perf. 1 p.p. Pres. Perf. 3 p.s. Pres.																	1		1	 1
FUNZIONARE to work	3 p.s. Pres.													4	2						1
GIOCARE to play	1 p.s. Pres. 3 p.s. Pres. 3 p.s. Pres. Prog. 3 p.p. Pres. Prog Inf. 1 p.s. Pres. Perf. 1 p.p. Pres. Perf.									1			1	1 1 2 1	 2 1 1	 1	 1	 1 1	 2	 1	 1
GIRARE to turn	1 p.s. Pres. 3 p.s. Pres. 3 p.s. Pres. Prog. 2 p.s. Pres.												4 2	2		 1 1		 2			

GONFIARE to inflate	Inf.																		3			
GUARDARE to look	2 p.s. Imp. Aff. 3 p.s. Pres. Inf. 3 p.s. Pres. Prog.				33	5	7		1	1	5	9		4 2	3	4		1 1	1 5			5
LAVARE to wash	2 p.s. Pres.														1							
LEGGERE to read	3 p.s. Pres. 1 p.p. Pres. 3 p.s. Pres. Perf.				1						1				3 2							
MANCARE to miss	3 p.s. Pres.																					1
MANGIARE to eat	3 p.s. Pres. 2 p.s. Pres. 3 p.s. Pres. Prog. 1 p.s. Pres. 3 p.p. Pres. 3 p.s. Pres. Perf.			4	2		2	3	1		2				12 1 1	1 4		1	3 2	1 1 5		
METTERE to put	1 p.p. Pres. 1 p.s. Pres. Inf. 2 p.s. Imp. Aff. 3 p.s. Pres. Perf. 2 p.s. Pres. 1 p.p. Pres. Perf. 3 p.s. Imperf.									2		1			1	1 4 1 1 2	1 1 2 2		2 2	2 2 1	1 5 1 1 1	
MUOVERSI to move	3 p.s. Pres.																		4			
NASCONDERE to hide	3 p.s. Pres. Perf.																					3
PARLARE to speak	2 p.s. Pres.									1												
PENSARE to think	1 p.s. Pres.																					1
PERDERE to lose	1 p.s. Pres. Perf. 3 p.s. Pres. Perf. 1 p.p. Pres. Perf.																1			1 2		1

PESCARE to fish	3 p.s. Pres. Prog.																1				
PIACERE to like	3 p.s. Pres.														3						
PIANGERE to cry	3 p.s. Pres. 3 p.s. Pres. Prog.							1							 1						
PICCHIARE to beat	Inf.																		1		
PIOVERE to rain																					
PORTARE to bring	3 p.s. Pres. 1 p.s. pres. Perf.						1						2		1 1		1	 1			
POTERE can	3 p.s. Pres. 1 p.p. Pres. 1 p.s. Pres. 2 p.s. Pres. 3 p.p. Pres.															9 7			 1	 1 2 2	 7
PRENDERE to take	Inf. 1 p.s. Pres. 3 p.s. Pres. 1 p.s. Pres. Perf. 2 p.s. Pres. 1 p.p. Pres. 2 p.s. Imp. Neg. 1 p.p. Pres. Perf. 1 p.s. Pres. Perf. 3 p.s. Imperf. 3 p.p. Imperf.				3 3		1					1 1		1	 3 1			5 1 1 1		2 2	1 1 1 2
PULIRE to clean	3 p.s. Pres. Prog.														1						
RACCONTARE to tell	1 p.s. Pres.																				3
RICORDARSI to remember	2 p.s. Pres.																			1	
RIFARE to do again	Inf.																				1
ROMPERE to break	3 p.s. Pres. Perf. 3 p.s. Pres.													1	2	1				1	2

SALTARE to jump	3 p.s. Pres.									4												
SAPERE to know	1 p.s. Pres.												5	6	1	6	8	9	2	9	4	
SCECLIERE to choose	1 p.s. Pres. Perf.																1					
SCENDERE to go down	3 p.p. Pres.																				3	
SCIVOLARE to slip	Bare P.P.													1								
SCRIVERE to write	3 p.s. Pres. Prog.																	1				
SEDERE to sit	2 p.s. Imp. Aff. 1 p.s. Pres.									1				2								
SENTIRE to hear	Inf.																				1	
SOFFIARE to blow	3 p.s. Pres. 3 p.s. Pres. Perf.															1 1						
SPARARE to shoot	3 p.p. Pres. Perf.																		1			
SPAVENTARE to frighten	2 p.s. Imp. Neg. 3 p.s. Imperf.																				3 1	
SPINGERE to push	1 p.s. Pres. 2 p.s. Pres. 1 p.s. Pres. Perf.														2	2 3 3						
STACCARE to detach	3 p.s. Pres. Perf.																1					
STARE to stay	2 p.s. Imp. Aff. 3 p.s Pres. Inf. 3 p.p. Pres.								1					1		3 1				3	1	
SUCCHIARE to suck	Inf.																	1				
SVEGLIARSI to wake pu	3 p.s. Pres. Perf.																1					