Bilingual First Language Acquisition: The Nature of the Weak Language and the Role of the Input

Francesca La Morgia
SALIS
Dublin City University

Supervisor: Prof. Eithne Guilfoyle

PhD 2011

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____
Student N. : 56126395
Date: _____

Acknowledgments

I would like to express my sincere gratitude to my supervisor Prof. Eithne Guilfoyle. Her advice, inspiring attitude and way of thinking have been of great value for me. I am also grateful to Prof. Antonella Sorace for the insightful discussions and constructive advice during my stay at the University of Edinburgh. I wish to thank my colleagues at the University of Edinburgh who made my visiting time an incredible learning experience.

I gratefully acknowledge the financial support of the Irish Research Council for the Humanities and Social Sciences.

This work would not be accomplished without the support of my family and my friends: my deepest and more sincere gratitude is for them. Finally, I wish to sincerely thank all the families and children who participated in this research.

Table of Contents

Abstract

Bilingual First Language Acquisition:
The Nature of the Weak Language and the Role of the Input

Francesca La Morgia

This thesis investigates the development of the weak language in early bilingual language acquisition and its results are based on longitudinal and experimental data from 4 Italian-English bilingual children and their parents.

The purpose of this thesis is twofold: firstly, to present a new method to assess weak language development and the role of the input in bilingual first language acquisition; secondly, to determine whether there is a relationship between input, weak language development and the acquisition of new information structure.

The factors included in the analysis of the weak language are rate of acquisition, production of target-deviant forms, vocabulary, MLU and discourse pragmatics. The results are summarised in the Weak Language Scale. The results are further tested by examining longitudinal and experimental data which are used to test the hypothesis that children who develop Italian as a weak language have difficulty processing subject inversion structures, which require a high processing load due to the interface between syntax and pragmatics.

The results of the Weak Language Scale are then compared to those of the Input Scale, which represents the amount of qualitative and quantitative input each child has been exposed to.

The final results show that the input plays a major role in bilingual first language acquisition and it has an effect on weak language development. The findings also suggest that linguistic properties at the interface between syntax and pragmatics are harder to process for children who develop Italian as a weak language.

List of tables

List of figures

CHAPTER 1
Introduction and theoretical assumptions

1.1 Introduction

This thesis investigates the nature of the weak language and proposes a new method to assess language dominance in young bilingual children who acquire two languages simultaneously from birth. The initial assumption underlying this study is that balance or equal development of the two languages is infrequently attested in simultaneous bilingual first language acquisition (2L1 acquisition). Children acquire a strong and a weak language, which can show different patterns in all linguistic domains[1]. The first studies that analysed the differences between the strong and the weak language were carried out in the 1990s (Schlyter 1993, Genesee, Nicoladis and Paradis 1995, Lanza 1992, 1997), and this topic has recently become the subject of much research in the area of bilingual language acquisition (Meisel 2007, Müller and Pillunat 2008, Cantone et al. 2008, Bonnesen 2009). As many of these studies have demonstrated, the strong language develops similarly to a first language (L1) in monolinguals, while the weak language is somewhat different. However, it has not yet been fully discovered in what ways the weak language differs from the strong one and how these differences can be reliably tested.

The main issue that needs to be addressed in this regard concerns methodology: the studies carried out so far have analysed a variety of factors, examining different sets of data, without producing a unified method of assessment of the weak language. Therefore, the first aim of this thesis is to identify and test some features which can be considered markers of weak language development (also on the basis of findings from previous studies), and ultimately to propose a new method of analysis. Once the criterion to assess the weak language is established,

[1] Studies on the weak language have interested all linguistic domains (see chapter 2). However, in this thesis I will focus on syntax, morphology and lexicon.

the second important phase of the analysis concerns the causes of weak development. Recent studies have highlighted the possibility of a relationship between input and dominance, but so far no study has provided evidence in favour of this hypothesis (Rothman 2009, Bonnesen 2009).

To summarise, the main aims of this study are to establish a method to assess the weak language and to find out whether there is a relationship between input and weak language development. The following research questions will be addressed:

- What are the characteristics of weak linguistic development?
- Do properties at the interface between syntax and pragmatics represent a difficulty for children developing Italian as a weak language?
- What is the relationship between the input and weak linguistic development?

In order to answer these questions, I employ spontaneous and experimental data as well as a questionnaire which investigates each child's linguistic background. The corpus of spontaneous longitudinal data was collected by audio-recording four bilingual children and their Italian parent(s) over a period of one year. The parents' data is used in the analysis of the input, while the children's data is used in that of the weak language. The parents also completed the *Questionnaire on the linguistic background of the bilingual child* (Appendix B and Appendix C), which was used to gather information on the amount of exposure to Italian at home (chapter 6).

In order to test the production of subject inversion (chapter 5), two elicitation tasks were designed and administered to the four children, as well as control groups[2].

1.2 UG and the role of the input

The framework adopted to answer the research questions presented in the previous section is the generative theory of Universal Grammar (UG), which claims that humans are born with an innate structured linguistic knowledge (Chomsky 1975, 1981). Every human is endowed with this universal language faculty from birth. UG provides rules that apply to all languages and it guides the acquisition of language-specific parameters. Therefore UG is made of principles, which are universal rules, and parameters, which vary depending on the language. The UG theory explains not only the structure of any natural language, but also the process of language acquisition, which is based on the combined action of the innate system and evidence

[2] A full discussion of the data is provided in chapters 3 and 5.

coming from the external environment. According to this theory, children do not need formal instruction or correction in order to learn to talk, and positive input is sufficient to allow them to set the language-specific parameters. The grammar is an abstract entity which allows children to make use of external evidence in order to acquire language. The existence of an innate internal system explains how, although they are exposed to meagre evidence, children can acquire very complex languages and construct sentences that they have never heard before in a relatively short amount of time (Chomsky 1981, Cecchetto and Rizzi 2000, Vallauri 2004). This theory, known as the *poverty of the stimulus argument* (Chomsky 1981), states that the evidence children are exposed to is insufficient on its own to account for the complexity of the language acquisition process and for the ability to produce potentially infinite combinations of words. However, this argument does not fully account for the role of the input in 2L1 acquisition. The hypothesis formulated in this thesis is that the linguistic imbalance which is commonly found in bilinguals can be explained by examining the exposure to the input and that there is a relationship between the quality and quantity of input and weak language development.

No study to my knowledge has yet provided a detailed analysis of the relationship between weak development and the quality and quantity of the input. This issue will be addressed in chapter 6, which focuses on the analysis of quantitative and qualitative aspects of the input and highlights the relationship between input and weak language development.

1.3 Further theoretical assumptions

The generative framework is also adopted for the analysis of the syntactic structures of Italian. The syntactic analysis presented in this thesis focuses on word order, and in particular the position of subject (S) and verb (V). In Italian, subjects can appear before the verb (preverbal) or after (postverbal), as shown in the following examples:

1.1 Gianni è arrivato (SV)
Gianni has arrived

1.2 È arrivato Gianni (VS)
*Has arrived Gianni
Gianni has arrived

Subject distribution is governed by syntactic and pragmatic rules (see chapter 2 – section 2.5 and 5). The pragmatic constraint governing the position the subject is the information load of the sentence: if the subject represents old information, it appears in the preverbal position (1.1); if it represents new information it appears in postverbal position (1.2). A further constraint concerns the thematic structure of verbs. Postverbal subjects can occur with any type of verb[3], as shown in the following examples with unaccusative (1.3), unergative (1.4), and transitive verbs (1.5), however, they are more commonly found with unaccusative verbs, because of their argument structure (see chapter 5).

1.3 È arrivato Gianni
*Has arrived Gianni
Gianni has arrived

1.4 Ha telefonato Gianni
*Has phoned Gianni
Gianni has phoned

1.5 L'ha fatto Gianni
*It has done Gianni
Gianni did it

As I will show in chapter 5, postverbal subjects are used significantly less than preverbal subjects and overall are very infrequent in the spontaneous data from the four bilingual children. For this reason, I will test their production in two elicitation tasks, in order to determine whether the scarce occurrence of postverbal subjects is determined by processing difficulty caused by the complexity of the structure, which can be attributed to the simultaneous activation of syntactic and pragmatic knowledge[4].

Another element of Italian syntax which is discussed in this thesis (chapter 4) is the distribution of overt and null subjects. While in some languages the subject has to be overt, in others, such as Italian, it is possible to omit it.

[3] For a more detailed analysis of the types of "inversion verbs" see Pinto (1997). An early formulation of theories on subject inversion can be found in Belletti (1988, 2001).

[4] Since this thesis is mainly concerned with the syntax-pragmatics interface in relation to the acquisition of the weak language, I will not focus on the constraints related to different verb types.

1.6 È arrivato
**Has arrived*

1.7 Gianni è arrivato
Gianni has arrived

The occurrence of null subjects, which is common to other languages such as Spanish and Catalan, has been attributed to the fact that these are morphologically rich languages[5]. It also has to be considered that the choice between null and overt subject is governed by discourse conditions (see chapter 4, section 4.9). Pro-drop is a parameter, therefore children have to set it according to the language acquired. It has also been found that subjectless sentences are commonly produced by children who are acquiring a non-pro-drop language. An explanation for this phenomenon is that there is a default parametric value that makes children produce null subjects until they are exposed to sufficient evidence to set the parameter appropriately (Hyams 1986). More recent analyses have demonstrated that there are differences between early null subjects produced by children who speak pro-drop languages and those produced by children who speak non-pro-drop languages[6]. Children who acquire Italian set the pro-drop parameter very early (around age 2), and it has been shown (Rizzi 1994, Guasti 2000) that their subject omission occurs in the same contexts as the adult language. On the basis of this claim, it is possible to assume that children develop knowledge of verbal agreement and they are therefore able to produce null and overt subjects in the appropriate contexts (Guasti 2004). Following this hypothesis, I will assume that the bilingual children who participate in this study have correctly set the pro-drop parameter. However, as other studies have shown (Serratrice, Sorace and Paoli 2008), children acquiring Italian and English (respectively a pro-drop and a non-pro drop language) could fail to select the appropriate option in Italian, and they may produce more overt subjects than monolingual Italian children. The phenomena of subject omission and subject inversion are both used in this thesis to assess weak language development. The starting assumption is that (as well as other factors), failure in the selection of null/overt subject and in the production of postverbal subjects in the appropriate pragmatic contexts can be seen as a sign of linguistic weakness. This hypothesis will be discussed in more detail in chapter 4 and 5.

[5] Null subjects are found also in other languages such as Chinese and Japanese. However, this type of null subjects underly different syntactic phenomena (Jaeggly and Safir 1989).
[6] See Guasti (2004) for a more detailed review of these studies.

1.4 Organisation of the thesis and chapter outline

The thesis is organised as follows:

Chapter 2 presents a theoretical discussion of the weak language, the acquisition of structures at the interface between syntax and pragmatics and the role of the input, based on a review of previous studies and on the research questions addressed in this thesis.

Chapter 3 constitutes an introduction on the linguistic background of the four case studies of bilingual Italian-English children. In this chapter, I present the methodology of data collection, providing a descriptive analysis of the data and an overview of each child's linguistic background.

In chapter 4, I propose an analysis of the weak language, examining the children's rate of acquisition, production of target-deviant forms, MLU, lexicon and discourse pragmatics. The results are summarised in the Weak Language Scale.

Chapter 5 deals with the production of postverbal subjects in bilingual children. In this chapter, I examine longitudinal and experimental data in order to test the hypothesis that children who develop Italian as a weak language have difficulty processing subject inversion structures because they require a high processing load related to the interface between syntax and pragmatics.

Chapter 6 explores the role of the input. In this chapter, the quantity and quality of the input are analysed by examining the spontaneous data from the bilingual children's parents and the results of the questionnaire on the child's linguistic background. The results are summarised in the Input Scale and compared to those from the Weak Language Scale.

In Chapter 7 I draw the final conclusions, evaluate the results achieved, discuss the contribution to knowledge, and identify areas for further research.

CHAPTER 2

The study of the weak language

2.1 Introduction

The first issue addressed in this thesis concerns the identification of the characteristics of the weak language and the development of a method for their assessment. Many studies on child and adult bilingual development mention the distinction between weak and strong language, but in some cases this judgement is based on the perceived proficiency in the two languages rather than on a systematic analysis. Some recent research has examined different aspects of child language acquisition in order to identify some features that can be associated with weak language development and it has provided a valuable contribution to the understanding of the processes underlying 2L1 acquisition (see section 2.3). However, the main difficulty in the study of the weak language is the lack of common criteria of analysis and of a standardised methodology.

In this chapter, I evaluate some of the methods and results in the main studies that have addressed the issue of language dominance and I provide an overview of the factors they have analysed. In addition, I examine previous studies on the acquisition of subject inversion, in order to show how the analysis of this phenomenon could contribute to the understanding of weak language development.

Finally, I review previous research findings on the role of the input in language acquisition which support my initial hypothesis that there is a relationship between weak language development and exposure to the input.

The studies reviewed in this chapter constitute the background of my research and of the initial assumptions which have brought me to formulate the following research questions:

- What are the characteristics of weak linguistic development?

- Do properties at the interface between syntax and pragmatics represent a difficulty for children developing Italian as a weak language?
- Is the lack of input a cause of weak linguistic development?

2.2 Bilingualism terminology

The focus of this thesis is the development of the weak language in bilingual children who acquire two languages from birth. This type of simultaneous bilingualism is considered to be equivalent to the acquisition of two first languages and it is referred to as Bilingual First Language Acquisition (BFLA) or 2L1 acquisition[7]. In addition, a child is considered a simultaneous bilingual if the two languages are acquired before the age of 3 or 4 years, and early successive bilingual if they are acquired afterwards (Unsworth 2005). We can observe two early maturational phases in child language acquisition: the first one takes place before the age of four, and if both languages are acquired before this age, they are considered to be two L1s. The second phase takes place after age four, when the grammar of the L1 has mostly been acquired (Guasti 2004, Meisel 2004). If the second language is introduced during this second phase of maturation, occurring between the age of four and puberty, the child still has the potential to achieve native-like competence in both languages, but as he/she grows up, the acquisition process becomes increasingly less spontaneous (Unsworth 2008, Meisel 2008, Rothweiler 2008). The third phase takes place after puberty. Several studies have explored the nature of the so called critical period from different perspectives and most of the results demonstrate that after this stage it is difficult (if not impossible) to attain native-like mastery of the L2 (Birdsong 1999, Long 1990). As I will show in this chapter, age is not the only factor affecting bilingual first language development, and simultaneous bilingual acquisition does not necessarily result in equal attainment in the two languages.

Another terminological distinction will be made in order to differentiate the two languages spoken by the children. The children analysed in this study are simultaneous bilinguals who acquired Italian and English from birth in Ireland, a predominantly English-speaking country. Since English is the language spoken by

[7] The terminology related to bilingualism that will be used in this thesis is based on Li Wei's (2000) classification (see Appendix A), as well as on current research on child bilingual acquisition (Meisel 2007, De Houwer 2009).

the majority of the population, it will be referred to as the majority language, while Italian will be referred to as the minority language. I wish to stress that this distinction does not constitute a prediction of dominance in the child's linguistic development. As I show in this chapter, the contexts in which the minority language is used are quite limited in comparison to those in which the majority language is spoken. A further distinction will be made between the weak and the strong language, in this case referring to the children's individual performance (see chapter 4 - section 4.2 for a more detailed discussion of terminology). I will therefore refer to minority or majority language when describing the status of the language in the external environment and to weak or strong when referring to the children's performance.

2.3 The weak language

In this section, I review some of the most relevant research findings on weak language development, focusing on the methodologies that have been employed. Throughout this thesis, I will use the terms *strong* and *weak* language, rather than *weaker/stronger* or *weaker/dominant*, which are generally used in other studies. (see chapter 4 – section 4.2)

Assuming that the majority of bilinguals manifest dominance in one of the two languages (Meisel 2007), it is necessary to clarify how this dominance emerges. It is also important to stress that, as the two languages develop, balance can shift, and each language can be weak or dominant to a varying degree at a particular time. The strong/weak dichotomy does not have to be seen as a weighing scale, in which the progress of one language is proportional to the failure in the other[8].

Most of the studies conducted so far have shown that the weak language is different in some aspects from the strong language, but they have not examined a sufficient number of factors to establish a method to assess imbalance across languages and at different stages of development. More recent research has highlighted the need for a more comprehensive and standardised method of analysis, based on the assessment of different groups of speakers (Cantone et al. 2008). However, to my knowledge, no study has so far provided a detailed

[8] In fact, research has demonstrated that it is possible for a child to have two equally strong or weak languages (Müller and Pillunat 2008).

description of the characteristics of weak language development also in relation to the analysis of its causes.

One of the first studies of the weak language was carried out by Schlyter (1993), who analysed the linguistic development of Swedish-French bilingual children living in Sweden. Her analysis shows that children who developed French as a weak language made errors in word order, finiteness and agreement, and they exhibited difficulty producing multi-word utterances. Comparing the data from children with French as a weak language and L2 learners of French, Schlyter found similarities in the types of errors and concluded that the weak language resembles an L2. This theory has since been revised by the author, and confuted by other researchers (Döpke 2000, Meisel 2007, Bonnesen 2009). The method used by Schlyter to identify a weak language was based on the analysis of MLU values and norm-deviant forms in the French production. According to the author, other indicators of weakness are the difficulty in using multi-word utterances where required and the occurrence of mixing stronger language structures into the weak language. As she observes, some of the errors affect word order and agreement:

> If the language is only slightly weaker, the child may use personal pronouns, but place them in an incorrect position, and /or combine them with a verb which is not marked for finiteness. He/she may use the correct verb form to mark past or future tense, but fail to mark person/number agreement correctly. The word order may be more incorrect than in a corresponding sample of the stronger language.
>
> (Schlyter, 1993: 296-297)

As a result, Schlyter claims that the weak language may exhibit errors of finiteness, word order and placement of negation, which are also common in the production of adult L2 learners of French.

Schlyter's method of analysis was replicated in another study (Bonnesen 2009), which examined the same syntactic structures comparing the norm-deviant forms produced by German-French bilingual children with French as a weak language with those produced by German L2 learners of French (table 2.1).

Table 2.1 Bonnesen's analysis of the weaker language

Norm-deviant forms in French L2 (Bonnesen 2009)
DP-subjects are hardly ever used with co-referent clitics
DP-subjects do not appear in right dislocated position
Subject clitics are used with both finite and non-finite verbs
Verbs are sometimes positioned between the subject clitic and the verb
There is no reduction of je before a vowel (je=1st person subject pronoun)
Low subject omission rate at early stages
Errors in the word order of the negation marker and the verb tend to occur

His results, which contradict Schlyter's, show that L2 learners produce errors that are not found in bilingual children with French as a weak language. The criteria used by Bonnesen to determine whether the children had French as a weak language were based on evidence showing that the two children (Françoise and Christophe) had a lower MLU in German than in French, they made several errors that are not common in L1 speakers, and they used French less frequently than German. Bonnesen considered another factor, namely the rate of acquisition, which he measured by analysing the increase in the number of verbs produced by the children. The results of this analysis revealed that the weak language follows the same acquisition patterns as the L1, but it is characterised by a slow development and by the production of errors which are not found in the L1.

The hypothesis of the similarity between the L2 and the weak language is confuted also in another study based on the analysis of English-German bilingual data (Döpke 2000). Döpke observes that it is possible that speakers of a weak language produce norm-deviant forms which are similar to the cross-linguistic influence errors made by L2 learners. However, she claims that the two types of linguistic development are fundamentally different and the emergence of structures in the weak language which may seem to indicate influence from the strong language (for example word order errors) should merely be attributed to processing difficulties that children may have when they have to select between two competing structures. The examples below show the types of word order

errors made by the German-English bilingual children (examples from Döpke 1998: 567-571). In sentence 2.1 the verb *essen* should follow the complement, in sentence 2.2 the finite verb *kommt* should precede the negator.

2.1 *Ich möchte essen das
I want eat that
'I want to eat that'
(Target word order: Ich möchte das essen)

2.2 *Hund nicht kommt rein
dog not come in
'(the) dog doesn't come in'
(Target word order: Hund kommt nicht rein)

The assessment of the weak language proposed by Döpke is based on the observation of word order errors like those shown in the examples above and on the comparison of MLU values. According to her analysis, the three bilingual children produce shorter sentences in English than in German and they use English word order in German sentences but do not use German word order in English ones. Short MLU and processing difficulties are therefore the two main indicators of weak language development.

The three studies reviewed so far base their distinction between weak and strong language on the analysis of MLU values and of the production of target-deviant forms. These two factors have proven to be useful in detecting differences between two sets of data, and they will be included in my analysis of the weak language (see chapter 4).

Another factor that has been analysed to test dominance is code-switching. Different studies showed evidence supporting the hypothesis that language dominance influences the directionality of code-switching (Petersen 1988, Genesee, Nicoladis and Paradis 1995, Lanza 1992, 1997, Bernardini and Schlyter 2004). According to Lanza, children with a weak language mix functional categories from the strong into the weak language. This theory is also supported by evidence presented by Bernardini and Schlyter (2004), who propose the *Ivy Hypothesis*, which predicts that unbalanced bilingual children are likely to project syntactic structures from the strong into the weak language, which 'grows like ivy on the structural tree of the Stronger Language' (p. 49). Their analysis of the weak language is based on MLU and Upper Bound (length of the longest

grammatically structured utterance in a transcription), but it does not focus on the characteristics of weak language development.

The hypothesis of a relationship between code-switching and dominance has recently been challenged by studies which have provided evidence of the independence of the two phenomena. These studies have shown that code-switching can be triggered by social and situational factors and it can occur independently of language dominance (Cantone 2007, Cantone et al. 2008). It also has to be taken into account that not all children use code-switching, and even those who have one very underdeveloped language might not use it at all. It is possible to argue that imbalance can determine the production of mixed utterances, in cases where the child needs to compensate the lack of knowledge of a word or structure. However, it would be difficult to verify in each case which are the factors responsible for this phenomenon. Since there is not yet agreement on whether the directionality of code-switching can be a useful indicator of weak language development, I analyse the mixed utterances produced by the children and I will determine whether this factor should be included in the final analysis.

The different studies reviewed in the previous sections show that there is not yet a unified methodology for assessing dominance. The main problems are the selection of data for comparison, the lack of normative data, and the difficulty in selecting the appropriate factors to analyse and test their significance. The factors that have been included in the analysis of the weak language are MLU, Upper Bound, MMU (average percentage of multi-morphemic utterances), code-switching directionality, lexical acquisition and production of target-deviant forms (Genesee, Nicoladis and Paradis 1995, Schlyter 1993, Cantone et al. 2008, Bonnesen 2009). By analysing these factors we might find out that one language develops faster, has a richer lexicon or is used more frequently and with more fluency than the other. The studies carried out so far show some of the differences between the strong and the weak language, but they do not provide a comprehensive and reliable method of assessment and do not account for the different degrees of weakness. In my analysis, I will follow the methodological suggestions proposed by Meisel:

> The first consists of examining the grammatical development of children who use "unusual" constructions […] in one language, which might qualify as the weaker one. The second is to tackle the problem from a grammar-external perspective, analyzing the language use of children whose development in one language is delayed or who tend to avoid using one of their languages, searching for developmental patterns or constructions that are typically not found in the language of the respective monolingual or of balanced bilingual children.
>
> (Meisel 2007: 500)

Another fundamental methodological issue in the study of the weak language is the choice of data. While in some studies the children have been compared to adult L2 learners or to other bilinguals, it has emerged that it is important to create a norm on the basis of which it would be possible to assess language development in different languages and across different groups of speakers.

Two studies provide an important contribution to the development of methodologies in this area. The first one (Arias et al. 2005) draws a comparison between the two languages in bilingual children and also between monolinguals and bilinguals. By looking at different groups of speakers and different languages, Arias et al. demonstrated that there is some degree of variation among different languages in the acquisition of some linguistic domains. Therefore, they argue that if we find delay in a child's linguistic development, we should try to determine whether this phenomenon is common among other monolingual and bilingual children who speak the same language. For this reason, they argue that the analysis of dominance in bilingual children requires also a comparison with monolingual data (also see Rolla San Francisco et al. 2006).

The second study (Cantone et al. 2008) proposes an original approach to the data selection, with the aim to establish a norm for comparison. The authors suggest that the most comprehensive methodology should be based on determining the monolingual norm for each language and then comparing bilingual to monolingual development and establish a bilingual norm. This type of analysis partly resolves the issue of comparing heterogeneous groups and should avoid the problem of basing bilingual analysis only on a monolingual norm. To create a norm, the authors analyse the French data from monolingual and bilingual children (French-German) and the German data from monolingual and bilingual children (Italian-German).

The types of bilingual development they predict to find are:

A Both languages develop in a normal fashion

B Language1 develops in a normal fashion
- language2 develops faster than the norm
- language2 develops slower than the norm

C Both languages deviate from the norm
- both languages are higher than the norm
- both languages are lower than the norm
- language1 is higher, language2 is lower than the norm

(Cantone et al 2008: 323)

According to the authors, data supporting type A and C are less frequent. As they suggest, their method for assessing the differences between two languages could be criticised for two reasons: the first is that it is not possible to determine whether the case studies analysed can be considered representative of other similar bilingual children. The second is that there is no clear distinction between the dependent and independent variables that are included in the analysis (MLU, Upper Bound, number of utterances per recording session and increase of noun types). The authors finally conclude by suggesting that describing language development in terms of dominance and distance from the norm does not account for bilingual variation:

> Studying the distance between languages in bilinguals merely shows us which language might develop faster, but this does not imply that it "dominates" the other language. […] Only if we compare the children's development to some "bilingual norm" we may be able to tell whether a language which develops more slowly than the other is also "weak".
>
> (Cantone et al. 2008: 337)

The methodologies used in Cantone et al. (2008) and Arias et al. (2005) provide the basis for the choice of data in this thesis (see chapter 4).

To summarise, studies on unbalanced development mainly reached the conclusion that the development of the weaker language differs at least in some respects from that of the stronger language, of an L1 and of an L2. Different

phenomena have been found to be typical of a weaker language. These have been analysed separately and also in different combinatorial sets, but no study has provided a convincing description of the characteristics of weak language development, both in terms of acquisition milestones at the early stages and of prediction of attainment at a later stage. It also has to be noted that the diversification in the methods of analysis make it difficult to determine which results are more reliable. Table 2.2 enumerates the phenomena that have been associated with weak language development.

Table 2.2 Characteristics of the weak language development

WEAK LANGUAGE DEVELOPMENT
1. Slow rate of acquisition of syntactic structures or functional categories
2. Production of target-deviant forms
3. Limited vocabulary
4. Limited/no switching into the weak language
5. Frequent switching from the weak into the strong language
6. MLU consistently shorter than L1 children
7. MMU consistently shorter than L1 children
8. Infrequent initiation of conversation in the weak language
9. Avoidance of complex structures
10. Avoidance/refusal to use the weak language

Table 2.3 Factors analysed to assess the weak language

Factors	Linguistic Domains
MLU	Syntax-Lexicon-Morphology
Upper bound	Syntax-Lexicon-Morphology
MMU	Syntax-Lexicon-Morphology
Direction of mixing	Syntax-Lexicon
Lexical variety/size	Lexicon
Word types	Lexicon
Verb types and tokens	Lexicon
Number of utterances per recording session	Syntax-Lexicon
Emergence of functional morphemes	Morphology
Percentage of correct consonants	Phonology
PLMU (Phonological MLU)	Phonology

Table 2.3 presents the different factors analysed in studies on the weak language and the corresponding linguistic domains they belong to. All the factors enumerated in the tables above have been so far considered equally significant, so there is not yet an indication of what phenomena are more relevant for the study of the weak language. In my analysis, I select MLU, age, rate of acquisition, production of target-deviant forms, vocabulary, and discourse pragmatics (the last factor has not been previously analysed in studies on the weak language). These factors have been chosen because they are typically used in the assessment of language development and also because it was possible to find monolingual and bilingual data for comparison (see chapter 4). The major difficulty is to prove the reliability of the data used for comparing different groups of children. As Cantone et al. (2008) have demonstrated, to have a comprehensive type of analysis, it is necessary to create a norm on the basis of other bilingual as well as monolingual data. In chapter 4, I employ both monolingual and bilingual data from previous studies to provide a more comprehensive overview of the development of Italian (Serratrice 1999, Bernardini 2004, Cipriani et al. 1993,

Antelmi 1997, Lorusso, Caprin and Guasti 2004, Ferrari and Matteini 2008, Guasti et al. 2008). As it has been found in the previous studies of the weak language, there are limitations to this type of analysis. The first limitation lies in the availability of the data: in order to create a reliable norm for the development of Italian in bilingual children, it would be necessary to analyse a larger sample in order to account for individual variation. The second limitation is the reliability of the choice and analysis of factors used to assess weakness. A further discussion of these issues will be provided in chapter 4.

2.4 Word order and subject inversion

The second question addressed in this thesis concerns the acquisition of properties at the interface between syntax and pragmatics. The starting hypothesis is that children who develop Italian as a weak language have difficulty processing postverbal subjects. As I explained in the outline of the theoretical framework in Chapter 1, the order of constituents in Italian is subject to syntactic and pragmatic constraints and it determines the interpretation of the sentence. When the subject of the sentence represents new information it appears after the verb. This Verb-Subject (VS) structure is called subject inversion, because it consists in the inversion of the position of the subject from preverbal to postverbal. This inversion requires an operation on a syntactic level, which is driven by the interpretation of the sentence. Therefore, the processing of subject inversion requires the activation of syntactic and pragmatic knowledge, an operation that requires a high processing load, and that may therefore constitute difficulty for non-native speakers - even for the near-native ones - (Belletti, Bennati and Sorace 2007) and for some bilingual children (Müller 2008, Hinzelin 2003). My initial assumption, based on some of the findings from the studies I review in this section, is that children who develop Italian as a weak language will have difficulties processing interface structures which are governed by syntactic and pragmatic constraints.

In recent years, linguistic research is increasingly focusing on the interplay between different domains, both from theoretical and developmental perspectives (Avrutin 1999, Burkhardt 2005, Hopp 2007, Lozano 2009, Sorace and Serratrice 2009, Wilson 2009). These studies have demonstrated that many language properties cannot be fully explained by examining a single linguistic domain, but

they are the result of the interplay between two domains. Research on the acquisition of subjects in different languages provides evidence showing that the analysis of interfaces is necessary for explaining the functioning of many linguistic structures. As many studies on L1 acquisition demonstrate, postverbal subjects are acquired early, and they occur mostly with unaccusative verbs. However, evidence from 2L1 acquisition shows that some bilingual children have difficulty mastering this structure. According to Adragão and Costa (2003), the strong evidence from the postverbal position of subjects with unaccusatives demonstrates that there is an unmarked position for the arguments, which allows the child to be aware that the VS order is an available option. More evidence of the acquisition of subject inversion comes from a study on the acquisition of Romanian. Avram and Coene (2003) found that Romanian children use postverbal subjects significantly less frequently than adults. Their longitudinal data of two children aged 1 to 2 years shows that postverbal subjects are very limited in the spontaneous production (respectively 12% and 5.9%) and they conclude that early postverbal subjects are not adult-like, since adults use postverbal subjects in more contexts and more frequently. This finding is supported by Grinstead (2004), who claims that at the very early stages of language development (before age 2) children do not have access to the knowledge of pragmatic distinction between new and old information. However, Adragão and Costa (2003) present evidence showing that children are able to distinguish between old and new information very early, therefore the lack of postverbal subjects in the data should not be attributed to "late mastery" of subject inversion (see also Kapetangianni 2007, 2008).

The available Italian data on the production of postverbal subjects in monolingual children shows that they acquire subject inversion early and they use this structure productively (Lorusso, Caprin and Guasti 2004). However, as it has been demonstrated in studies on 2L1 acquisition, bilingual children might have difficulty achieving this competence (Hinzelin 2003, Müller 2008). This difficulty could be due to the fact that the processing load of an interface structure such as subject inversion may be too high for some bilinguals. Following Adragão and Costa (2003) and Lorusso, Caprin and Guasti (2004), I assume that postverbal subjects are acquired early both by monolingual and bilingual children. However, I aim to test the hypothesis that some bilingual children have difficulty processing subject inversion structures and this difficulty can be related to weak language

development. Further evidence supporting my theory is provided by studies of adult L2 acquisition of postverbal subjects (mostly based on data from speakers of a non-null subject language who acquire a null-subject language). It has been demonstrated that learners exhibit difficulty in producing subject inversion structures when these are not available in their native language (Belletti and Leonini 2004). Belletti, Bennati and Sorace (2007) also tested the production of postverbal subjects by native English speakers on the basis of four off-line tasks, which included videos, story-telling, picture verification and a headline task[9]. Their results demonstrate that adult L2 learners (even near-native ones) use postverbal subjects in a non-native like fashion. The explanations that have been provided to justify the difficulty in the mastery of subject inversion are the complexity of the structure (Belletti, Bennati and Sorace 2007) and also the input, as Rothman explains:

> Pragmatic conditions are imparted to language learners on the basis of discourse patterns. Logically, for a learner, child and adult alike, to acquire [these] pragmatic features, the discourse pattern that they are exposed to must provide unambiguous evidence as far as this is concerned. It is possible that a contributing factor to the delay in acquiring subject distribution in L2 Spanish and/or any 'residual optionality' for some highly advanced learners has to do with the input they receive.
>
> (Rothman 2009: 968)

On the basis of these findings, I assume that testing the production of postverbal subjects in bilingual children could lead to interesting results which could provide a further insight into the understanding of weak language development.
Having established that subject inversion requires a high processing load, which makes the structure hard to acquire for L2 learners and for some bilinguals (but not for monolinguals), I intend to test the hypothesis that this processing difficulty is found only in bilingual children who develop Italian as a weak language.

[9]In the headline task the participants had to reconstruct the order of a sentence that constituted the headline of a newspaper article.

2.5 The role of the input

The last question addressed in this thesis concerns the relationship between input and weak language development. Once we have determined how to assess a weak language, it is important to start examining the possible causes[10] of this type of development. The main hypothesis is that the amount of input in the two languages affects 2L1 development. To test this hypothesis, I will analyse the quality and quantity of input the children are exposed to by examining the parents' data and by using a questionnaire.

Research on the input has shown that children are sensitive to sounds from the earliest days of life. It has also been demonstrated that they can soon distinguish between different languages (Oller et al. 1997, Bosch and Sebastián-Gallés 1997, 2001). Especially in the early phases of life, children are primarily in contact with their parents (generally their mothers) and siblings, and, depending on the situation, to a varying extent, they are exposed to the external environment. As many studies have shown, the main source of input is represented by the individuals who interact with the children (Vigil, Hodges and Klee 2005; Pancsofar and Vernon-Feagans 2006; Pan et al. 2005). These are mostly parents[11], at least at the earliest stages of life, in particular before the child starts attending a day-care centre or school. The importance of parental input has been stressed since the 1970's, when researchers found that mothers use a specific register when addressing their children. This register, called *motherese*, has features that make it different from the speech used among adults and is characterised by restricted vocabulary, short and simple sentences, slow and repetitive speech, high pitch and exaggerated intonation (Snow 1972; Gleitman, Newport and Gleitman 1984; O'Grady, 1997). The *motherese hypothesis*, as proposed by Gleitman, Newport and Gleitman (1984), suggests that this form of speech is necessary for the child in order to acquire language. The main evidence against the *motherese hypothesis* comes from studies on mother-child discourse in several different countries, which show that not all parents use a special register to address children, and in some countries children are rarely involved in conversation with adults until they learn to talk (Stoll and Lieven 2008, Jurugo

[10] Different linguistic and non-linguistic factors can account for weak language development. In my analysis, I will only consider the role of the input.

[11] In the category of parents, any equivalent sort of carer/guardian is implied. This applies to any general reference to parents throughout this thesis.

2009). Even though the *motherese hypothesis* is now considered obsolete, studies on child directed speech (CDS) have demonstrated that parental input has an impact on the child's lexical development. For instance, it has been shown that a child's first words are often those that the parents use more frequently and that the size of the child's vocabulary is directly proportional to the amount of vocabulary used by the parents. Studies of the relationship between the amount of productive vocabulary and input provided in the conversation between mother and child (Hoff and Naigles 2002) reveal that parents' input is beneficial for both lexical and syntactic development. Hoff and Naigles (2002) also suggest that the process of acquisition is the result of the combination of cognitive factors and the understanding of the context. The role of the interlocutor's engagement in conversation is also important for the child in order to understand the meaning of words (Akhtar and Tomasello 2000, Baldwin 2000). Therefore, the non-linguistic context, together with the lexical content and syntactic structure of the input, constitute the information that children use to acquire word meaning (Hoff and Naigles 2002). These studies mostly show that exposure to the input may impact on lexical acquisition. However, they do not make claims related to grammatical development.

More evidence regarding the importance of the input is provided by studies on specific language impairment, language delay and deprivation. Windsor, Glaze and Koga (2007) found that children who were raised in a severely deprived environment showed less developed language skills compared to their peers living in the same communities. The results show major differences between children raised in orphanages and their peers raised by their biological family in the scores of MLU, lexical and phonological development. In another study on the interaction between mother and child, Vigil, Hodges and Klee (2007) found a relationship between the mother's input and the children's linguistic delay. The mothers of children with language delay were found to give minimal feedback to their children, providing responses that were not appropriate in the specific context and not related semantically to the child's utterance. Even though there is no consistent evidence of a direct relationship between the quality of parental input and the child's linguistic delay, this study shows that children with parents who actively engage in conversation and show responsiveness produce a higher number of utterances earlier than those whose parents fail to do so. Extreme cases

of feral children who were completely deprived of linguistic input also represent evidence of the importance of exposure to language[12]. The famous case of Genie (Curtiss 1977) shows that the rare interaction with other speakers and the extremely limited input can result in the failure of linguistic development.

All the studies on the role of the input show that the environment in which the child is raised is an important factor in language development, especially when children acquire two or more languages[13] (Unsworth 2007). Research on 2L1 acquisition has demonstrated that some bilingual children are less exposed to each language than monolinguals, while others are constantly exposed to an almost equal amount of input in the two languages in a variety of contexts (Nicoladis 2008, Paradis 2008). However, even studies on multiple language acquisition in bilingual environments have shown that balanced bilingualism is not the norm and that the child develops one language more than the other, at least in some areas (generally the lexicon). An interesting study on the role of the input and bilingual language use (De Houwer 2007) bases its results on a questionnaire which aims to find out why some children exposed to two languages from early on "fail" to master one of the two languages. About 2000 families completed a questionnaire on language use. In all the families that participated, at least one of the parents spoke Dutch, the majority language. The results of the questionnaires revealed that all children succeed in acquiring the majority language, while the minority language appears to be less widely used. This seems to be related to several factors, but mainly to parental language input. De Houwer suggests that the reason for the "failure" is the more limited input in the minority language, but she does not explore in depth the role of input frequency.

[12] It has to be taken into account that feral children also experience emotional trauma and non-linguistic factors have to be considered.

[13] The analysis of the role of the input in bilingual acquisition often involves the investigation of both languages. This might not always be possible, since different problems may arise. For example, in countries where the majority of the population is bilingual, the two languages are likely to be spoken both at home and in the environment. In these cases it may be hard to determine in which language the input is more quantitatively and qualitatively consistent. In order to get a full picture of the amount of input and output the child receives in the two languages it is necessary to be in constant contact with the child, and to be aware of the linguistic environment inside and outside the home. Moreover, children are exposed to the two languages in different contexts (especially children growing up in predominantly monolingual environments), with different people, and the balance of the input in each language can shift frequently.

Balanced (or almost balanced) bilingualism does not only depend on parental input, but also on other factors such as the language spoken by the majority and the status each language has outside the child's home. Among the case studies presented in this thesis (chapter 3), there are three children who speak the minority language with only one parent, while they use the majority language with the other parent, his or her family and also with other children. The fourth child uses the minority language at home with both parents and the other language only when she plays with children in the neighbourhood. Overall, the four children use Italian only when interacting with the parents and rarely when visiting family and other Italian-speaking families. Therefore, the parents represent for these children the main source of Italian input. Since the use of Italian in the family can be easily quantifiable, I employ the *Questionnaire on the linguistic background of the bilingual child* (Appendix B and Appendix C) in order to explore some aspects of the children's linguistic environment. Similar bilingual studies based on questionnaires demonstrated the effectiveness of this method in providing an overview of the child's linguistic background (Sorace et al. 2009, Paradis 2008, Nicoladis 2008, De Houwer 2007). In this thesis (chapter 6), the questionnaire is not only used to obtain information on the family language use, but it is also a source of data for the quantitative analysis of the input. The questionnaire[14] is completed by the parents at the beginning and at the end of the research period in order to keep a record of the characteristics and the changes in the child's linguistic environment, and to provide a background to the data collection. It includes questions regarding the languages used at home and in the child's environment, the amount of time spent using each language, and the family's attitude towards Italian.

The input received from the parents and the external environment has so far been presented as the main factor affecting 2L1 acquisition. The following table shows some of the factors that emerge from different accounts in the literature on bilingualism. I will consider some of these factors in more detail in chapter 3 and 6. As I will explain in chapter 6, some of the factors outlined below, such as attitude, are more difficult to describe and measure. For this reason, I will mainly focus on measurable linguistic aspects.

[14] The questionnaire is in English and in Italian (see Appendix B and C). Parents are asked to complete it together, where possible.

Table 2.4 Factors analysed in the literature

Input	Amount of time the child is exposed to the two languages (quantity of input) / Richness and variation in vocabulary and syntax (quality of input)
Language use	Contexts in which the child uses or is exposed to each language
Child's attitude	Child's behaviour and attitude in respect to the use of the two languages (enjoyment – play - frustration – refusal)
Parents' attitude	Parents' behaviour, attitude and strategies in relation to the child's use of the two languages (language choice - encouragement – indifference – disregard – disinterest)

Ultimately, the aim of the analysis of the input is to determine whether there is a relationship between this factor and weak language development. This hypothesis has been put forward in previous studies, which claim that the amount of exposure to a language is directly related to the "success" in acquisition (Schlyter 1993). By stating that the majority language will be the stronger language, Schlyter assumes that children are likely to perform better in the majority language, presumably because that is the language they are exposed to more often. Argyri and Sorace (2007) also point out that among the factors that cause unbalanced bilingualism, the amount of exposure to the input can be considered a determining one. The same argument is supported by Grosjean, who argues that 'the main reason for dominance in one language is that the child has had greater exposure to it and needs it more to communicate with people in the immediate environment' (1982: 189). Other researchers have pointed out that the child's ability in using the two languages and the development of the lexicon proportionally increase with exposure to the input (Döpke 1992, De Houwer, 1995; Hoff and Naigles 2002). As these findings show, the input can be considered a determining factor in the development of two L1s, especially in contexts where one language is spoken by a limited number of individuals in the child's environment. Although many studies highlight the importance of the input, none of the ones focusing on the weak language analyse the children's

interaction with their parents and their overall exposure to the two languages. Moreover, there is not yet to my knowledge a study that examines quantitative and qualitative aspects of the input in relation to weak language development.

Table 2.5 and 2.6 provide a summary of some recent studies of the role of the input in L1 acquisition with a brief description of the methods employed both in monolingual and bilingual studies.

Table 2.5 Methods employed to measure the input in L1 acquisition

Study	Age Groups	Language/s	Method
Lidz, Gleitman and Gleitman (2003)	24 children (3;2 - 3;10)	Kannada	Experiment with transitive/intransitive verbs
Theakston et al. (2004)	9 children (1;10 – 2;11)	English	Measurement of input frequency (mother's utterances) / child's acquisition of verbs
Pancsofar and Vernon-Feagans (2006)	120 children (1 – 3)	English	Analysis of language input of mothers and fathers measured in terms of output, vocabulary, complexity, questions, and pragmatics during triadic free play sessions
Huttenlocher et al. (2002)	34 children (4 – 5)	English	Analysis of CHILDES data to compare syntactic complexity in mother and child's utterances
Borovsky and Elman (2006)			Experiments using computational simulations
Valian and Casey (2003)	29 children (2;6 – 3;2)	English	Spontaneous speech and experiment with speaking puppet
Westergaard (2004)	3 children (1;9 – 3)	Norwegian	Analysis of topicalization constructions in the data
Wijnen et al. (2001)	2 children from the CHILDES database	Dutch	Measurement of input frequency (based on mother's utterances) in relation to child's production of root infinitives
Vigil et al. (2005)	60 children (2)	English	Comparison of input from parents of children with and without language delay
Behrens (2006)	1 child (1;11 – 4;11)	German	Measurement of input frequency and quality (mother's and father's utterances)
Hoff and Naigles (2002)	63 children (1;6 – 2;4)	English	Analysis of maternal and child's speech based on number of utterances, word tokens, word types, MLU
Strömqvist and Richthoff (1999)	2 children	Swedish	Measurement of input frequency (mother's and father's utterances) and feedback morphemes

Table 2.6 Methods employed to measure the input in bilingual acquisition

Study	Age Groups	Languages	Method
De Houwer (2007)	4,500 children and their families	Dutch and other additional Languages	Questionnaire on language use
Paradis and Navarro (2003)	3 children (CHILDES)	English and Spanish	Comparison of production of null and overt subjects in children and parents
Nicoladis (1998)	1 child (1;0 – 1;6)	Brazilian Portuguese and American English	Calculation of productive vocabulary based on recordings and parents' notes

This selection of studies on the role of the input in child language acquisition shows that a frequently employed method for assessing the quality and quantity of the input is the analysis of speech samples in the parents' and children's language. This method is generally employed both in monolingual and bilingual acquisition. However, there are few bilingual studies that focus on the role of the input, and they mostly deal with the comparison between the two languages.

By analysing a small sample of bilingual children, I aim to look closely at the quantity and quality of the input, combining the use of a questionnaire with the analysis of spontaneous longitudinal data. After examining the amount of input each child is exposed to, I compare these results to those from the analysis of the weak language, in order to determine whether the exposure to the input can account for the differences between the children who develop Italian as a weak or a strong language.

2.6 Chapter summary and conclusion

In this chapter I have introduced some of the theories on the development of the weak language, the processing of structures at the interface between syntax and pragmatics and the role of the input. It emerges from studies on weak language development that there is not yet a unified methodology to assess dominance, since several factors and linguistic domains have been analysed in different ways. Another important element to be taken into consideration is the type of data employed to analyse language development in young simultaneous bilinguals and to create a "norm" on the basis of which it is possible to compare different groups of children.

The linguistic domains which are more commonly analysed to assess dominance are lexicon, syntax and morphology. Some of the factors that many studies have analysed, such as MLU, lexical development and production of target-deviant forms have been shown to account for some of the differences between weak and strong development and some of them will be included in my analysis (chapter 4). In addition, I wish to test two factors that have not been previously included in weak language studies, namely distribution of overt and null subjects and subject inversion. The choice of these two factors is based on the finding that subject inversion and selection of null vs. overt subjects may constitute difficulty for some bilingual children.

Finally, in this chapter I examined some results from studies on the input in monolingual and bilingual language acquisition to provide ground for my hypothesis that the quality and quantity of input the child is exposed to can determine whether a child develops a strong or weak language.

From the review of the literature, the following main findings have emerged:

- **More research is necessary to understand the nature of the weak language**

The overview of the studies on the weak language has shown that weakness has been investigated from different perspectives and by using various types of analyses. Overall, the weak language has been found to develop differently from the dominant language in bilinguals and from an L1 in monolinguals.

- **Subject inversion constitutes processing difficulty for some bilinguals**

Subject inversion is a structure which requires the activation of syntactic and pragmatic knowledge, because the speaker needs to be aware of the information structure of the sentence and accordingly move the subject to a postverbal position. The findings from the studies on the acquisition of postverbal subjects show that they are acquired as early as preverbal subjects by monolingual children. However, it emerges from studies of bilingual children and L2 adult learners that the interplay of syntax and pragmatics can represent difficulty due to the complexity of the structure and the simultaneous activation of two domains.

- **The input plays an important role in 2L1 acquisition**

The quantity and quality of input and the use of the languages in the child's home and in the environment have been found to affect different aspects of linguistic

development and also the dominance of one language over the other. Therefore, the input should be investigated to determine whether it has an impact on the development of a language as weak or strong.

The results from the studies reviewed in this chapter suggest that the three areas investigated can be studied in combination in order to find the relationship between input and weak language development and to determine whether subject inversion can be tested to assess the weak or strong development of Italian.

CHAPTER 3
Overview of the longitudinal data

3.1 Introduction

This chapter provides an overview of the spontaneous longitudinal data collected for this study, the methodology of data collection. For each child, I describe the duration and frequency of the recording sessions, the language use in each recording considering child's age, percentage of unintelligible, Italian, English and mixed utterances and MLU. Finally, I provide examples of the interaction between the children and the Italian interlocutors (the mother or the investigator) at different stages, focusing on different phases of language development. The linguistic development of the four bilingual case studies is compared to that of other monolingual children from one of the CHILDES corpora (Cipriani et al. 1993) and other bilingual children from two studies of Italian-English (Serratrice 1999) and Italian-Swedish children (Bernardini 2004). These studies are briefly described in section 3.3.7, since their data will be employed for comparative purposes in the analysis of the weak language (chapter 4). As previous studies have demonstrated (Cantone et al. 2008), the choice of data to use for studying language dominance is a determining factor in the reliability of its assessment. As I show in this chapter, the four Italian-English bilingual children constitute a rather homogeneous sample, which is the main requirement for a study on weak language development. While previous studies have assessed dominance by examining data from children which they assumed were weak in a given language (Bonnesen 2009), in this thesis I start from the assumption that any of the children could be developing Italian as a strong or weak language.

3.2 Development of Italian as a minority language

Language development follows specific patterns and stages across all languages. Children initially use one word at a time, generally to request or indicate the existence of objects and their dynamics, to describe actions, to ask questions, and to attribute properties to objects (Tomasello and Brooks 1999). These types of one word utterances constitute the first phase of linguistic development, which generally occurs when the child is about 12 months old (Guasti 2004, Keren-Portnoy et al. 2008). At around 20 months children start combining individual words, but their utterances lack some functional categories. Between age 2 and 3 children produce the first utterances containing functional categories, use infinitive verbs in main clauses, omit subjects and over-regularise tenses of irregular verbs (Radford 1990, Caselli, Casadio and Bates 1999). These phenomena recur across languages almost at the same developmental stages (Guasti 2004). The general L1 acquisition phases can also be applied to bilingual acquisition. In terms of syntactic acquisition, given sufficient input, bilingual children can achieve the normal L1 developmental milestones at the same time as monolingual children. However, as studies on the weak language have demonstrated, some bilingual children develop one of the two languages at a slow rate and they produce more errors than monolingual children. The most comprehensive method to analyse the child's linguistic development is a longitudinal study. While there are some longitudinal studies of the acquisition of Italian as an L1 (Cipriani et al. 1993, Antelmi 1997), only a few have analysed the development of this language in minority contexts[15]. In this thesis, Italian is considered the minority language, since it is spoken by a small group of people in the child's environment. However, the concept of minority has no qualitative connotation. I therefore base my analysis of the data on the assumption that all the children involved in this study are native speakers of both languages and that they might develop Italian as a weak or strong language. I will use the term minority language following the definition proposed by Rothman for the concept of heritage language:

> Like all monolingual and childhood bilingual learners, heritage speakers are exposed naturalistically to the heritage language;

[15] See Serratrice (1999), Serratrice, Sorace and Paoli (2004), Müller and Hulk (2001), Müller et al. (2002), Cantone et al. (2008), Bernardini (2004), Kupisch (2007), Schmitz and Müller (2008).

> however, this language is by definition a nonhegemonic minority language within a majority-language environment. Since the heritage language is the family language used and heard in restricted environments, there are varying degrees of deterministic consequences for the complete acquisition and/or maintenance of the heritage language, depending on when and how the societal majority language is introduced.
>
> (Rothman, 2007: 360)

The children who participate in this study match the description provided by Rothman, since they are acquiring Italian in an English-speaking environment[16] and they are exposed to a relatively limited amount of input in the non-hegemonic language, especially outside their home.

3.3 Methodology

In the sections that follow, I describe the method used for selecting and recruiting participants and for collecting and transcribing the data. I also present an overview of the factors used to analyse the data and I introduce previous studies on the acquisition of Italian which will be used throughout the thesis for comparing the four bilingual children to other monolingual and bilingual Italian speakers.

3.3.1 Selection of participants

The bilingual children participating in this study were recruited on the basis of criteria that were established for this research. In order to be included in the study the children had to:

1. Be exposed to English and Italian from birth.
2. Live in Ireland for most of the year.
3. Be at the early stage of linguistic development.
4. Be able to produce at least one-word utterances in Italian.

The subjects taking part in the research are four bilingual children (three females and one male) that I have named Costanza, Francesca, Matelda and Paolo. Their age ranges from 1;11 to 3;1 at the time of first data collection session.

[16] None of the children have been exposed to the Irish language.

3.3.2 Recruitment of participants

The participants were recruited by placing a notice in an online magazine for Italians in Ireland and by contacting the Italian Playgroup, the Italian Embassy and the Italian Cultural Institute in Dublin. Before starting the recruitment process, I received approval from the Dublin City University Ethic Committee to conduct research with children and their parents. Before accepting to get involved in a longitudinal research on bilingualism, the parents were asked to read and sign a *Plain Language Statement* and an *Informed Consent Form*, which clearly explained the aims and scope of the research. Every effort was made to protect the parents' and their children's anonymity. Even though the research involved only one parent, the other one was made aware of the research purposes and the methodology and also signed the forms mentioned above. The documents were available in English and Italian to make sure both parents had full access to the information.

3.3.3 Recordings

The data was collected by audio-recording the spontaneous interaction between the children and their parents with a high-quality digital recorder. Each visit to the families lasted at least two hours, during which almost one hour was dedicated to the recording. I also kept a diary containing some information given by the parents, notes on the recordings and other relevant details and I encouraged the parents to keep notes on significant changes in their children's linguistic development. Only Italian speakers (generally the parent and the investigator, and on some occasions also a relative) were in the house with the child for the whole duration of the recording session.

3.3.4 Activities performed

During the recording sessions, the participants performed activities which are part of their daily routine. The children and their parents were involved in activities such as drawing, colouring, clay modelling, story-telling, object-naming, educational games as well as ball-games, jigsaws, role-playing with dolls and many more. Depending on the situation, the children played either with the mother or both with the mother and the investigator. At the initial stages, they

mostly played with the mother. Every effort was made to keep the situation as natural and spontaneous as possible.

3.3.5 Transcripts

After each recording session, the recordings were transcribed. All transcripts were later checked by a native speaker and the discrepancies were then double-checked.

3.3.6 The analysis of the data

The data collected is used to show the development of the minority language in the four bilingual children. In this chapter, I present a general overview of each child's background, followed by the content of the recordings, showing age, MLU, the percentage of Italian and English utterances and the percentage of mixed utterances, dividing them according to the directionality of mixing. The tables containing this information show the percentages of Italian, English and mixed utterances on a sample of 100 in each recording. The purpose of this classification is purely descriptive and it is useful to observe the patterns of language use at different stages. The introductory section on the child's linguistic background is then followed by a more detailed analysis of the one and two word stage and the comparison with other Italian data. The data employed for comparison purposes comes from studies on monolingual and bilingual language acquisition, as I show in section *3.3.7.*

3.3.7 Other available data

Several studies have analysed the development of Italian from a generative perspective, both in monolingual and bilingual research.

The most widely used monolingual Italian data set is available on the CHILDES database[17]. Only two bilingual corpora have so far been made available: the Italian-Dutch corpus is a picture description study of the production of subject ellipsis in children from 5 to 13. The Italian-German corpus contains 5 Italian and 5 German transcripts of data from a child who lives in Italy. The Italian data covers a period of about 4 months, from age 1;9.0 to 2;01.04 (Klammer 2006).

[17] The corpora from the Child Language Data Exchange System are available on http://childes.psy.cmu.edu/

While the monolingual Italian database is quite extensive, there is not an equivalent set of bilingual data. Nevertheless, there are two longitudinal studies which provide valuable sources of data: one is Serratice's (1999) study on the emergence of functional categories in a bilingual Italian-English child (between the age of 1;10.08 and 3;0.17). The other set of data comes from Bernardini's (2004) research on the acquisition of the DP in two bilingual Italian-Swedish children, Lina from 1;4 to 3;7 and Lukas from 2;0 to 3;7.4 years of age. Both Serratrice's and Bernardini's works present a longitudinal analysis of both languages. Serratrice's description of Carlo's daily routine shows that the child receives an almost equal amount of input in the two languages. Even though he lives in a predominantly English-speaking environment (in the UK), he is exposed to a substantial amount of Italian input from his Italian baby-sitter, who spends more than two hours with him 5 days a week on a regular basis. Serratrice observes that while in the nursery Carlo is exposed to English mostly in a polyadic fashion with only few sentences strictly addressed to him, at home he is exposed to dyadic exchange with his mother or with his babysitter, so he is more frequently exposed to Italian than English in one-to-one contexts. The MLU values show that Carlo has an almost constant development in both languages, but Italian seems to be his dominant language (Serratrice, Sorace and Paoli 2004). Lukas and Lina, the bilingual children studied by Bernardini (2004), show a different rate of acquisition in the two languages. Bernardini calls the strong language $L1_1$ and the weak one $L1_2$. Italian is the $L1_1$ for Lina and the $L1_2$ for Lukas. This distinction is based on the fact that the children mix more often when talking to speakers of the $L1_2$ and also that their MLU in the $L1_2$ is shorter independently from the language of the interlocutor (p. 42). Bernardini's and Serratrice's studies also refer to input and dominance, therefore it is interesting to compare their results to the ones obtained in this thesis.

The next section provides an overview of the spontaneous longitudinal data. The children's linguistic development is analysed on the basis of traditional methods, such as the calculation of MLU values, as well as an overview of the emergence of language with specific reference to some significant characteristics of each child's linguistic production.

3.4 Analysis of the spontaneous data: Francesca

3.4.1 Data collection

Francesca was 2;4.20 at the time of first recording. She was recorded for 13 months at monthly intervals. Information on the linguistic background was obtained by her parents, who filled out the *Questionnaire on the Linguistic Background of the Bilingual Child*. Francesca was always audio-recorded in her home in the presence of her mother and the investigator. The recordings always took place in the afternoon, which is the time of the day the child spends with the mother, speaking only Italian. Francesca generally spends the afternoon mostly playing alone or with her mother. The recordings lasted about 30-40 minutes at the beginning and then they were almost 1 hour long when Francesca started producing more complex utterances. During each session, several activities were performed (see section 3.5.3). Table 3.1 shows the child's patterns of language use. Francesca produces mostly Italian utterances when addressed in Italian, and her use of English decreases after age 2;6. After this stage, English is mostly used to compensate for the lack of vocabulary. Table 3.1 shows the age of the child, the total percentage of utterances (Tot), unintelligible sentences (U), Italian sentences (Italian), English sentences (English), Italian-English mixed sentences (It-En), English-Italian mixed sentences (En-It)[18] and the MLU value calculated in words. The aim of this classification is to show the child's pattern of language use across time and the trend of the mean length of utterance.

[18] English-Italian and Italian-English refer to the directionality of code-switching.

Table 3.1 Francesca's data

Age	U %	Italian %	English %	It-En %	En-It %	MLU
2;4.20	4.7	67.7	22	5.5	0	1.39
2;5.10	3.2	61.2	22.5	4.8	8	1.26
2;6.19	8.7	71.4	10.9	2.1	6.5	2.06
2;7.28	0	98.5	0	1.5	0	2.21
2;9.07	2.9	93.3	1.4	1.4	0.7	2.11
2;10.17	6.3	90	0.9	1.8	0.9	2.21
3;0.17	2.8	85.8	7.8	3.5	0	2.45
3;1.17	4.4	92.4	2.2	0.8	0	2.80
3;2.27	0.3	97.9	1	0.6	0	3.13
3;5.0	1.9	98	0	0	0	2.84

Figure 3.1 Overview of Francesca's data

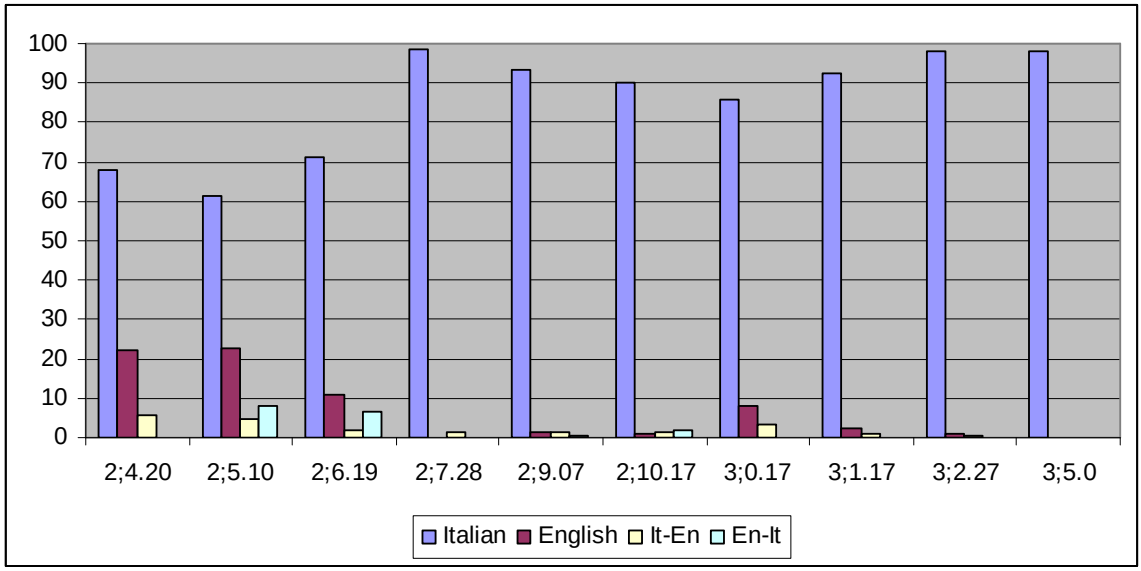


Figure 3.2 Francesca's MLU

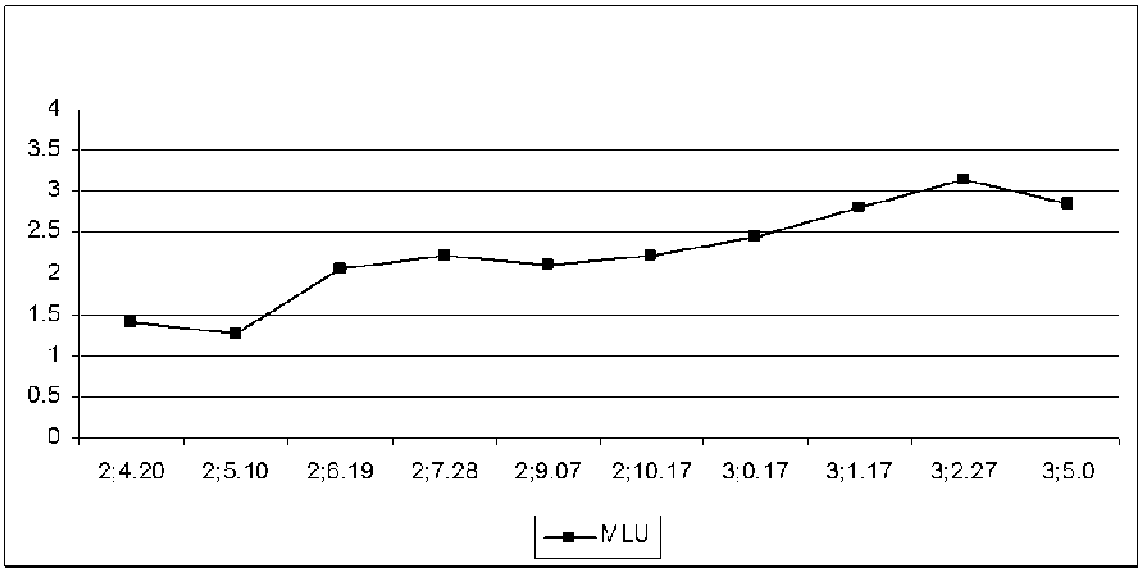


Figure 3.1 shows that Francesca initially produces mostly Italian utterances, but also many English and mixed utterances. As she gets closer to age 3, her production of Italian utterances increases. Her MLU shows an almost constant increase.

3.4.2 Francesca's linguistic development

Up to the age of 2;6 Francesca produces mostly one-word utterances, which are generally the repetition of her mother's last word (3.1, 3.2). She also produces a number of English and mixed utterances (3.3, 3.4). Her mother responds to these types of utterances by reformulating them in Italian. When she does so, Francesca repeats the reformulated sentence (3.5, 3.6, 3.7).

3.1 *MOT[19]:andiamo su?
shall we go upstairs?[20]
*FRA: su
upstairs

3.2 *MOT: questo libro non lo puoi leggere tu. È per grandi.
you can't read this book. It's for grown-ups

[19] The abbreviation used is based on the CHILDES system of abbreviation. *MOT= Mother; *INV= investigator, *Three letters of child's name (*FRA, *COS, etc.).
[20] Note on the translation: all examples in Italian have been translated into English. However, notes on the grammatical structure are only given where relevant.

*FRA: grandi
grown-ups

3.3 *MOT: questa è l'oca!
this is the duck
*FRA: duck! F[21] 2;4.20

3.4 *FRA: open questo.
open this F 2;5.10

3.5 *MOT: sì, eccola
yes, here she is
*FRA: eccola
here she is
*FRA: go downstairs
*MOT: andiamo giù
let's go downstairs
*FRA: diamo zu! (andiamo giù)
let's go downstairs

3.6 *MOT: quello non si puó prendere, vedi che è attaccato?
this can't be taken, do you see that it is attached?
*FRA: these are trousers
*MOT: eh, i pantaloncini corti.
yeah, shorts
*FRA: paccini (pantaloncini) F 2;4.20

3.7 *FRA: <u>open</u>[22] questo
open this
*MOT: come si dice?..a..
what do you say? ..o..
*FRA: <u>can you open</u> questo?
can you open this?
*MOT: come si dice?
what do you say?
*MOT: apri!
open!
*FRA: apri, mamma! F 2;5.10
open, mum!

[21] When reporting samples of children's speech, I will use the first letter of their name as an abbreviation.
The age of the children is written at the end of the last example referring to the age.
[22] The parts of utterances in English are underlined to highlight the point where the switch occurs.

Francesca's English utterances[23] are longer and more complex than the Italian ones and they all contain verbs, which are not yet produced in Italian (3.8, 3.9, 3.10). Initially, Francesca produces mostly English verbs.

3.8 *FRA: I'm going to sit down
3.9 *FRA: I'm watching telly F 2;4.20

3.10 *INV: ora che facciamo?
what are we going to do now?
*FRA: make this! F 2;6.19

After age 2;6 Francesca starts to consistently produce two-word utterances, consisting mainly of noun phrases (3.11).

3.11 *FRA: sporchi tutti.
all dirty F 2;6.19

There are no inflected verbs in the child's production up to the age of 2;7. Before the emergence of inflection, the only verbs are produced in the imperative form: *fai* (*do*), *guarda* (*look*), *apri* (*open*), *lascia* (*leave*), *siediti* (*sit down*). The first inflected verbs used productively by Francesca are the third person form of *essere* (*to be*) and *avere* (*to have*). She also starts using the verb *volere* (*want*) more and more frequently after age 2;9. However, she constantly uses the second person form *vuoi* instead of the first *voglio* and sometimes also instead of the third (3.12-3.17). This error is very frequent in Francesca's data, and it is documented up until the last recording, mostly with the verb *volere* (*want*). Francesca's mother often corrects her use of this verb, but the child continues to use the second person[24]. I hypothesise that this use of the second person over the first is strictly related to the input. Mothers often use the verb *want*, and, especially when addressing their child, they are likely to use phrases such as *'do you want x'* very frequently. While in English there is no difference in inflection between *I want* and *you want*, in Italian *volere* does not only require inflection, but it is also an irregular verb. Francesca's mother frequently uses *vuoi* when addressing her child and no instances of *voglio* are found in the mother's speech.

[23] A systematic study of the children's English was not conducted, therefore these observations are based on the English utterances that emerge in the Italian recordings.
[24] Francesca also uses *vuoi* appropriately for second person marking.

Therefore, the misuse of first person verb is probably a fossilised error that could be attributed to frequency of input.

3.12 *FRA: mamma, vuoi l'acqua ancora!
mum, (you) want water again! (I want water again!)
*FRA: vuoi acqua ancora!
(you) want water again!
*MOT: che vuoi?
what do you want?
*FRA: acqua
water F 2;9.07

3.13 *FRA: vuoi colorare
you want to colour (I want to colour) F 2;10.17

3.14 *INV: mi sa che mamma non vuole
I think mum doesn't want (you to watch TV)
*FRA: si, vuoi
yes, you want (she wants)
*INV: guarda, e' arrivata
look, she has arrived
*FRA: lo vuoi!
you want it (I want it)
*FRA: vuoi televisione!
you want TV! (I want TV)
*INV: ma quando c'è il sole non si guarda la televisione!
but when it; sunny you don't watch TV! F 3;0.17

3.15 *INV: e va a mangiare i pesci nel mare, guarda!
and it goes to eat fish in the sea, look!

*FRA: perchè vuoi mangi i pesci?
why do you want to eat fish?(why does it want…) F 3;1.17

3.16 *FRA: io vuoi togliermi scarpe.
you (I) want to take off my shoes. F 3;2.27

3.17 *FRA: vuoi yoghurt!
you want yoghurt! (I want yoghurt) F 3;5.0

From age 2;9 we can observe an increase in the use of longer sentences and inflected verbs.

3.18 *FRA: che c'è dentro questa?
what is inside this?

3.19 *FRA: do' (dove) sta casa lupo?
where is the wolf's house?

3.20 *FRA: che fai, ti metti giacca?
what are you doing, are you putting on (your) jacket? F 2;9.7

There is little evidence of deviant word order in Francesca's data. The few examples that can be found are in the use of the genitive.

3.21 *FRA: questo is Francesca.
this is Francesca's. F 2;7.28

3.22 *FRA: Marta macchina questa.
this is Marta's car F 2;9.7

3.23. *FRA: questo bambino's letto
this (is the) child's bed F 2;10.17

However, this deviant word order is found together with the correct word order.

3.24 *FRA: do sta passeggino Marta?
where is Marta's pram? F 2;9.7

After turning 3, Francesca produces longer sentences and a more varied lexicon. She makes only few gender and number marking errors, but she still has difficulties with verbs and articles (3.25-3.30).

3.25 *FRA: eh, non lo so
oh, I don't know
*FRA: mamma casa
mum house
*FRA: me con mamma fare spesa
me with mum go shopping
*FRA: mamma ha portato a crèche
mum has brought to crèche

Example 3.25 shows the use of sentences without verbs (*mamma casa*), the use of an accusative pronoun instead of nominative (*me* instead of *io*) and the use of an infinitive where an inflected verb should be used. The last sentence is missing a pre-verbal pronoun (*mamma <u>mi</u> ha portato*) and also the article preceding the word *crèche*.

3.26 *FRA: leggi questo
read this
*INV: questa è la storia degli amici?
is this the friends' story?
*FRA: tu leggi storia amici
you read friend story F 3;0.17

Example 3.26 shows two missing forms, an article (*la storia*) and a preposition (*degli amici*). As I will show in the following chapter, the omission of articles is a typical developmental phenomenon at the very early stage child language acquisition. Example 3.27 shows the use of an infinitive in place of an inflected verb, which is another typical developmental phenomenon (Rizzi 1994).

3.27 *FRA: dov'e' il porcellino?
where is the piglet?
*FRA: non si trovare il porcellino l'altro
can't be found the piglet the other
*FRA: non si trovare!
(it) *can't be found!*

3.28 *FRA: lei paura cagnolino, paura cagnolino
she fear doggy, fear doggy
*INV: nell'acqua, vedi, si vedono solo i piedi perché lui, splash!, è finito nell'acqua!
in the water, you see, you can only see the feet because he, splash! has ended up in the water!
*FRA: dov'è piedi?
where is feet?
*INV: eccoli
here they are
*FRA: vai dentro l'acqua piedi
you go into the water feet F 3;1.17

Example 3.28 shows the omission of the verb (*lei paura*, instead of *lei ha paura*) and of the preposition (*paura cagnolino* instead of *paura del cagnolino*) in the first sentence. The second sentence *dov'è piedi?* is a case of agreement error and again of article omission. The last sentence can be understood by looking at the context. Francesca is describing some pictures from a book, one showing a man that jumps in the water, the other one showing his feet emerging because his body is upside down. Francesca says "*go into the water feet*" while pretending to push the man's feet under the level of the water. So, the meaning of the sentence could be interpreted as "*he goes in the water with his feet*" or "*put the feet into*

the water" or "*the feet go into the water*". The sentence is not of easy interpretation mainly because of the lack of agreement.

3.29 *FRA: i salta
I jump

Sentence 3.29 shows a recurrent error in Francesca's use of the first person pronoun *io*. She often uses *i* (pronounced like the i of the IPA system) instead of *io*. This might be caused by the influence of the English first person pronoun. This sentence also shows a subject-verb agreement error similar to the ones shown in previous examples.

Sentence 3.30 shows a subject-verb agreement error (*io vai* instead of *io vado*) and also a number agreement error (singular compound preposition and plural noun).

Sentence 3.31 shows the use of the imperative instead of indicative form. The target form should be *io mi metto*, however this seems another error caused by the influence of the input. There is frequent recurrence of imperative forms in the mother's input. Therefore, we can assume that Francesca is using *metti+mi* instead of *mi+metto* because she hears the combination verb+pronoun more frequently. This is consistent with the observation that the first verbs Francesca produced are imperatives, probably because they are the most frequent in the input.

3.30 *FRA: io vai alla scimmiette
I go to the monkeys

3.31 *FRA: no, io mettimi scarpe
no, I put-me shoes F 3;2.27

To conclude, the data collected over 12 months show that Francesca's Italian has been constantly developing. This is demonstrated by her MLU and by the increasing complexity of her utterances. However, at 3;5 she still produces sentences that can be considered basic, mostly consisting of determiner, noun and verb, or verb and complement.

3.4.3 Comparison with other Italian data

Cipriani et al. (1993) divide the linguistic development in four different phases: the pre-syntactic phase (approximately occurring between 19 and 26 months), the primitive syntax phase (20-29 months), the nuclear phrase phase (24-33 months), and the phase of generalization of complex rules (27-36 months). If we use the classification proposed by Cipriani et al., we can observe that Francesca falls into the first phase until age 2;7 (31 months) at a stage in which the monolingual children are stepping to the third or fourth stage. Not only is the Italian children's MLU[25] on average longer, but they start constructing complex sentences earlier than their bilingual peers. Their utterances reach complexity (phase three) faster, and if we compare the last set of data collected from Francesca with any other monolingual child, we will see that her utterances are much less complex, resembling probably the nuclear phase. Overall, Francesca's development seems slower compared to monolinguals. This could be due to the lack of input, or to the reduced exposure to the minority language, which is typical of bilinguals. A further comparison with other bilingual children should help to cast light on this issue. Since the full corpora of bilingual children are not publicly available, the comparison is based on the analysis provided by the authors. The MLU can be considered a valid measure to compare children across different groups. In comparison to Francesca, Carlo has overall higher MLU values and constructs more complex structures at an earlier stage. Lina, the girl with Italian as a strong language, has an MLU pattern very similar to Francesca's, while Lukas, the child with Swedish as a strong language, has lower MLU values than Francesca's. Overall, it seems that, keeping age as a constant for comparison, bilingual children have lower MLU values. Compared to bilingual children, Francesca's Italian is weaker than Carlo's, but stronger than the "weakest" child Lukas.

3.5 Analysis of the spontaneous data: Costanza

3.5.1 Data collection

Costanza was 1;11.16 at the time of first recording. She was recorded every 5 or 6 weeks for 12 months. Information on the linguistic background was obtained by her parents, who filled out the "Questionnaire on the Linguistic Background of

[25] In all the studies on the development of Italian mentioned for comparison purposes, the MLU is calculated in words and not in morphemes.

the Bilingual Child". Costanza was always audio-recorded in her home in the presence of her mother and the investigator. The recordings take place on the day of the week in which the mother works only a half day. On those days Costanza is exposed to Italian for at least 4 hours. The recordings last between 20 and 40 minutes. During the sessions the child is involved in different activities, mainly involving games with toys and dolls. Table 3.2 shows Costanza's language use. Since the earliest stages, she shows separation of the two systems and she hardly ever uses English when addressed in Italian. Her vocabulary is varied from the earliest stages of linguistic development, and there are only very few instances of code-switching. There is also no evidence of deviant word order.

Table 3.2 Costanza's data

Age	Unintel. %	Italian %	English %	It-En %	En-It %	MLU
1;11.16	0	98	0	2	0	1.78
2;0.10	0	100	0	0	0	1.61
2;2.17	0	98.6	0	1.3	0	2.87
2;4.9	4.1	95.2	0	0	0.6	2.98
2;6.7	2.1	92.63	4.2	1	0	2.58
2;7.16	0	100	0	0	0	2.86
2;9.14	0	98.6	0	1.3	0	4.24

Figure 3.3 Overview of Costanza's data.

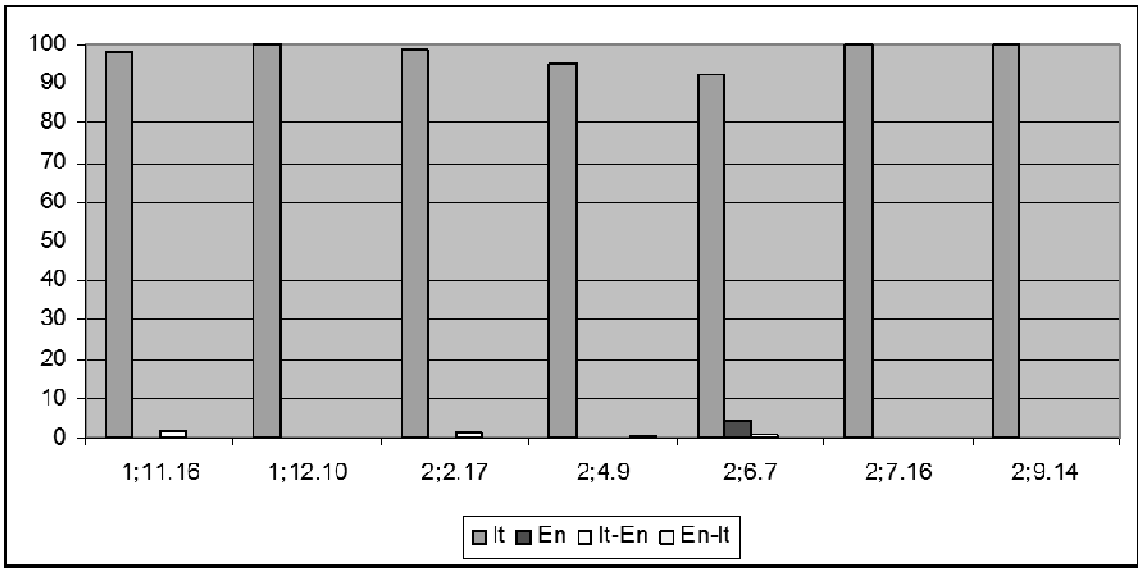
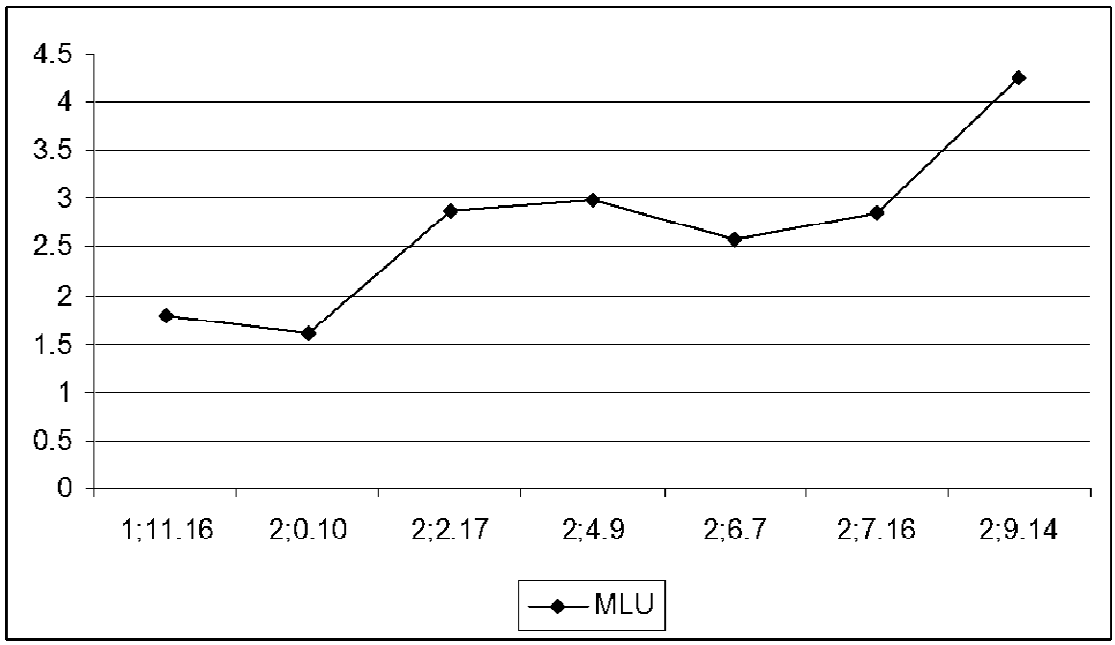


Figure 3.4 Costanza's MLU



Costanza's MLU follows mostly an ascendant trend and reaches a high value at a relatively early stage compared to the other three children of this study.

3.5.2 Costanza's linguistic development

The data collected during the first two recording sessions capture the early stage of word production. Costanza is the youngest child involved in this research. Since the first meeting, she has shown to have full understanding of every

sentence addressed to her, and she has always responded with no hesitation to everything she was asked (3.32). Unlike the other children, she seems to be very similar to a monolingual Italian child in terms of spontaneity in language production and the variety of lexicon used. Sentences 3.32-3.34 show samples of the interactions that took place during the recording sessions.

3.32 *INV: pronto, chi è? (pretending to answer the phone)
hello, who's this?
*COS: papi.
daddy
*INV: dov'è papi?
where's daddy?
*COS: a lavoro.
at work C 1;11.16

3.33 *INV: si mette un po' d'acqua
you put a bit of water
*COS: l'acqua dentro
the water inside
*INV: *me la cucini una salsiccetta?*
will you cook me a little sausage?
*COS: si
yes
*MOT: che gli fai?
what are you making her?
*COS: una pappa
some food
*MOT: vuole una salsiccia Franci
Franci wants a sausage
*INV: una salsiccia con un po' di patate, però, mica solo la salsiccia
a sausage with some potatoes, though, not just the sausage
*COS: no!
no!
*COS: non le patate, solo una salsiccia!
not potatoes, just a sausage!
*INV: una sola?
only one?
*COS: si
yes
*COS: no, no, casca la salsiccia!
no, no, the sausage is falling!

3.34 *COS: che fai, Francesca?
what are you doing, Francesca?
*INV: ti sto cucinando un po' di pasta, la vuoi?
I am cooking you a bit of pasta, do you want it?
*COS: no, brucia

no, it's too hot
*INV: brucia?
it's too hot?
*COS: si, però scotta
yes, but it burns
*INV: adesso la metto qua così si raffredda
now I put it here so it cools down
*COS: non scotta più?
it doesn't burn anymore?
*INV: no, l'ho messa fuori, vedi, non scotta
no, I have put it outside, you see, it doesn't burn
*COS: non scotta
it doesn't burn C 2;4.9

Costanza's early sentences show instances of article and verb omission, but correct gender and number marking. Unlike Francesca and Paolo, since the early stages she spontaneously starts a new topic in conversation, and she does not frequently repeat her mother's utterance. After the age of two, we can observe an increase in Costanza's production of new vocabulary, as well as the emergence of tense and of more complex structures. At this stage, Costanza masters gender and number marking, and the use of articles (3.35-3.36).

3.35 *COS: è bianco questo.
this is white
*MOT: e questo?
and this?
*COS: è blu.
it's blue.
*COS: eccolo il cane, eccola la nonna.
here is the dog, here is the granny
*INV: e tu di che colore ce li hai I calzini?
and what colour are your socks?
*COS: bianchi
white
*INV: ah, e ci sono i fiori rosa!
ah, and there are pink flowers!
*COS: rossi.
red

3.36 *INV: lo vado a mettere in cucina?
will I put it in the kitchen?
*COS: si
yes
*INV: dove esattamente?
where exactly?

*COS: qui, in questa cucina.
here, in this kitchen C 2;2.17

Costanza is the child with the highest average MLU. Moreover, her MLU is not only higher than bilingual children of the same age (Francesca and Matelda), but also higher than Paolo's, who is more than 12 months older than her.
Finally, we can observe that subordinate clauses emerge at around age 2;6 (3.37) and are used productively by age 2;9.

3.37 *INV: ti vuoi mettere due scarpe diverse?
do you want to wear two different shoes?
*COS: questa!
this!
*COS: mamma, di' a Francesca devo mette questa
mum, tell Francesca (that) I have to put on this
*MOT: no, mettiti la scarpa che scivoli
no, put on the shoe or you'll slip C 2;6.7

In the last session, which took place when Costanza was 2;9, we can observe the increase of the MLU and the use of an growing number of complex sentences.

3.38 *COS: Franci, vieni perchè ci nascondiamo!
Franci, come here because we hide!

3.39 *INV: lasciamo la bambolina qua?
we leave the dolly here?
*COS: no, perchè se lei piange non possiamo uscire.
no, because if she cries we can't go out. C 2;9.14

Finally, it is worth noting that Costanza often speaks Italian when playing alone[26]. This shows that her use of Italian is not only restricted to the interaction with her mother, but it is her natural linguistic choice in other contexts.

3.5.3 Comparison with other Italian data

Costanza's Italian develops similarly to the Italian children from Cipriani et al.'s study. Both the MLU and the syntactic complexity of her utterances resemble a monolingual Italian speaker's. As I show in chapter 4, Costanza receives the highest amount of qualitative input. Her mother's sentences are longer and more

[26] This fact is confirmed by the parents and it is also documented in the last recording session, during which Costanza goes to play by herself in her tent and talks to her toys in Italian.

complex from the early stages, and she does not use phrases typical of child-directed-speech or baby-talk as much as the other parents. It has often been argued that bilinguals can develop a more limited vocabulary in each of their languages, but that their lexicon is equal to monolinguals if we combine the lexical items from the two languages. However, this does not apply to syntactic development. Costanza is an example of how the syntactic development of a bilingual can equal a monolingual's, even in contexts in which the child is exposed to a relatively limited amount of input.

The comparison with the bilingual children in other studies shows that Costanza's Italian is more developed than Lina's and Lukas'. Costanza's MLU is on average slightly higher than Carlo's, but the data shows that overall the two children follow a similar MLU pattern. The comparison with other 2L1 data suggests that Costanza's linguistic development resembles the development of a balanced or dominant bilingual.

3.6 Analysis of the spontaneous data: Paolo

3.6.1 Data collection

Paolo was 3;1.27 at the time of first recording. He was recorded every 4 to 6 weeks for 11 months. Paolo was always audio-recorded in his home in the presence of his Italian mother and the Italian investigator[27]. The recordings take place always on the mother's day off, during which the child is exposed to Italian for at least 5 hours. Each session lasts at least 1 hour, during which the child is involved in different activities (see section 3.5.3). Table 3.3 shows Paolo's pattern of language use.

[27] Except on one occasion when the Italian grandmother is visiting.

Table 3.3 Paolo's data

Age	U %	It %	En %	It-En %	En-It %	MLU
3;1.27	0	87.7	3.4	2.7	6.1	2.11
3;3.23	1	89	3.1	2	4.7	1.81
3;4.25	0	80.3	4.5	6	9	1.77
3;7.10	0	92	0	4	4	1.66
3;10.19	0	100	0	0	0	3.02
3;11.17	3.5	96	0	0	0.5	2.65
4;0.29	0.5	99.4	0	0	0	2.87
4;1.28	0	100	0	0	0	2.21

Figure 3.5 Overview of Paolo's data.

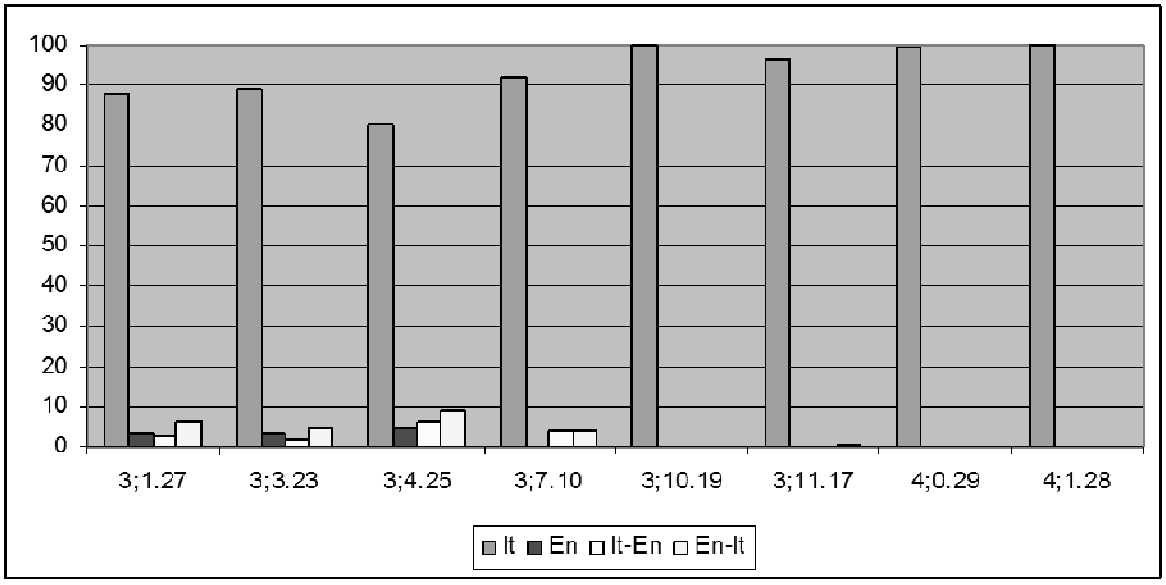
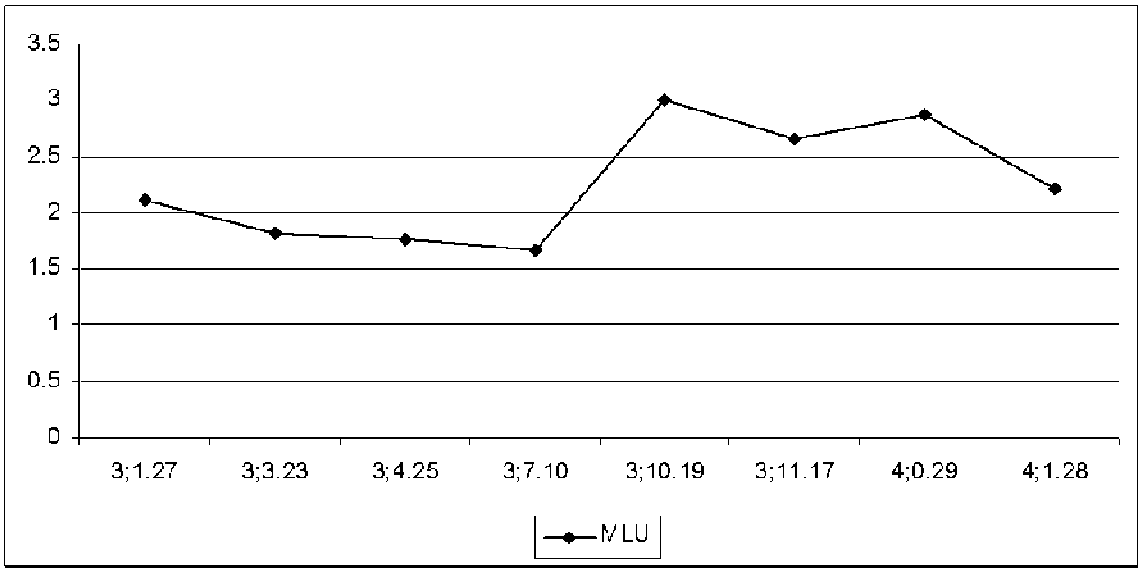


Figure 3.6 Paolo's MLU



Paolo uses mostly Italian, but up to age 3;10 he mixes in both languages, mostly from English to Italian. However, the last four recordings show an increase in the use of Italian utterances. His MLU has a fluctuating trend, especially if compared to Francesca and Costanza. Moreover, his MLU is on average shorter than that of children his age.

3.6.2 Paolo's linguistic development

Paolo is the oldest of the four children who participated in this study. His Italian is initially much less developed than his English, and also less developed in comparison to his bilingual peers. His Italian production at age 3;2 still shows some of the characteristics of the early stage of language acquisition. Initially most of his sentences are very short, and he produces incomplete words and utterances (3.40).

3.40 *INV: dove sta il lupo?
where is the wolf?
*PAO: là.
there
*INV: oh, è vero, non l'avevo visto!
oh, that's right, I didn't see it.

*PAO: sì, tivo[28].
yes, bad

[28] *tivo* stands for *cattivo*.

*INV: è cattivo.
it's bad
*PAO: tivo.
bad P 3;3.23

Paolo produces mixed utterances, but there doesn't seem to be a constant directionality. There are also cases of multiple switching within the same sentence (34). The switches occur mostly in the VP. While initially the mixed utterances are quite frequent, they gradually decrease at around age 4, when Paolo seems to master the Italian language with more confidence.

3.41 *PAO: guarda, he can't get out[29].
look, he can't get out.
*INV: non riesce a uscire.
(he) can't get out
*PAO: no, can't uscire là.
no, can't get out there.

3.42 *PAO: guarda!
look!
*INV: ah, una chitarra!
oh, a guitar!
*PAO: sì, I have one su.
yes, I have one upstairs

3.43 *MOT: dopo cena se sei buono.
after dinner if you are good.
*PAO: I'm buono
I'm good

3.44 *PAO: can't vedo un l'occo (un occhio)
I can't see an eye P 3;3.23

3.45 *MOT: e loro cosa fanno?
and what do they do?
*PAO: gone via
gone away (they go away)

3.46 *PAO: ma what's he got in mano?
but what's he got in his hand? P 3;4.25

Just before turning 4, Paolo starts producing more complex sentences. During this phase, he produces mostly interrogatives.

[29] The words underlined represent the switch to English.

3.47 *INV: ti piace questo libro?
do you like this book?
*PAO: si
yes
*PAO: perchè ** è un cigno?
*why ** is a swan?*
*PAO: but lui, lui…Il brutto anatroccolo è lui.
but he, he…The ugly duckling is him.
*INV: si, lui diventa un cigno.
yes, he becomes a swan
*PAO: e loro?
and they/them?
*INV: e loro anche
and they (do) too
*PAO: perchè lui diventa grande?
why does he get big/old?
*INV: perchè cresce
because he grows up
*PAO: perchè il cacciatore lui spara e prendono lui?
why does the hunter shoot him and they take him?
*INV: perchè il cacciatore spara agli uccelli per mangiarli. […]
because the hunter shoots birds to eat them […]
*PAO: perchè lui non ride?
why does he not smile?

At age 4 Paolo still omits determiners (definite articles more systematically than indefinite ones) and makes gender and number errors. Some of his utterances show an influence from English. In example 3.48, Paolo responds to his mother by using a typical English contrastive expression, *no, you are!*, which bears emphasis on the pronoun *you*. Italian doesn't have a similar construction, therefore this seems to be a case of cross-linguistic influence (*you are = tu sei*).

3.48 *MOT: stai facendo un po' di confusione, sai?
you are getting a little confused, you know?
*PAO: no, tu sei!
no, you are P 3;11.17

Unlike the other children, Paolo produces non-target word order that deviates not only from the English, but also from the Italian norm (3.51, 3.52). However, these cases are rare, and most of his deviant word order utterances seem to be modelled on the English word order (3.49, 3.50).

3.49 *PAO: lupo mano!
wolf hand (the wolf's hand)
TARGET: la mano del lupo P 3;3.23

3.50 *PAO: aranciona barca piccolo.
orange boat small
TARGET: barca arancione piccola/ piccola barca arancione P 3;11.17

3.51 *PAO: è l'acqua andato
(he) is the water gone (he is gone into the water)
TARGET: è andato in acqua/nell'acqua P 3;10.19

3.52 *PAO: cosa questo è?
what this is?
TARGET: cosa è questo? P 4;0.29

Overall, the development of Paolo's Italian seems to be slightly slower than the other bilingual children's. There also seems to be a delay in the production of structures that are generally acquired early in Italian. Paolo's sentences are often incomplete and his frequent mixing often makes his utterances hard to understand. He also seems to have difficulties in comprehension, and often asks his interlocutor to repeat. However, his data shows that his language is developing, even if at a slower pace.

The step to the production of multi-word utterances happens just before the age of 4. After this stage, Paolo produces more utterances and seems to be more confident in using Italian.

3.6.3 Comparison with other Italian data

Paolo's MLU values are lower than Carlo's. Even though the investigation of Carlo's development ends when the child is 3, we can observe that his MLU is on average 2.5 from the age of 2;5. The same values are found in Paolo's data only after the age of 3;10. Paolo's low and variable MLU values resemble Lukas', the child with Italian as a weak language. His MLU is also lower than Lukas' at some stages. What Paolo and Lukas have in common is mostly the lack of uniformity and consistency in the development. The data related to the input suggests that Paolo is likely to develop Italian as the weak language, and this is confirmed by this brief overview of the child's linguistic development, which will be further explored in the next chapter.

3.7 The analysis of the spontaneous data: Matelda

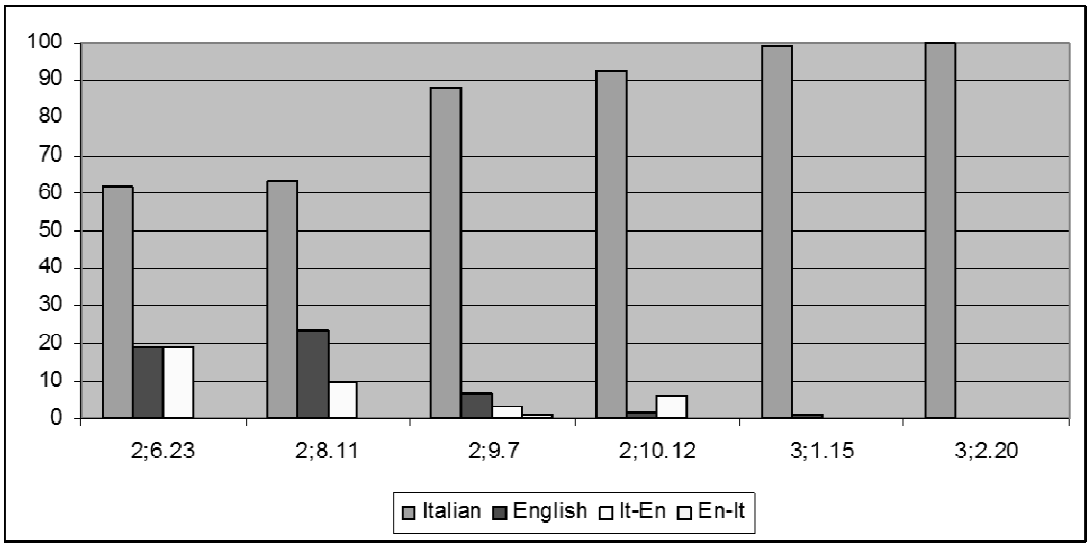
3.7.1 Data collection

Matelda was 2;6.23 at the time of first recording. She was audio-recorded for 9 months. The recording sessions are not as regular as the other children's because her family took two long breaks in Italy. Matelda is always audio-recorded in her home in the presence of her mother. During the meetings, the only people in the house are mostly the Italian mother and the Italian investigator, and sometimes an Italian nanny. The meetings last at least 1 hour and they take place generally in one room while performing different types of activities. Matelda is the only child with two Italian parents. She doesn't attend pre-school, and the environment she is raised in is predominantly Italian. As I have mentioned in the previous chapter, Matelda is the child who is exposed to the greatest amount of Italian input (over 80%). Table 3.4 and figure 3.6 show the child's patterns of language use.

Table 3.4 Matelda's data

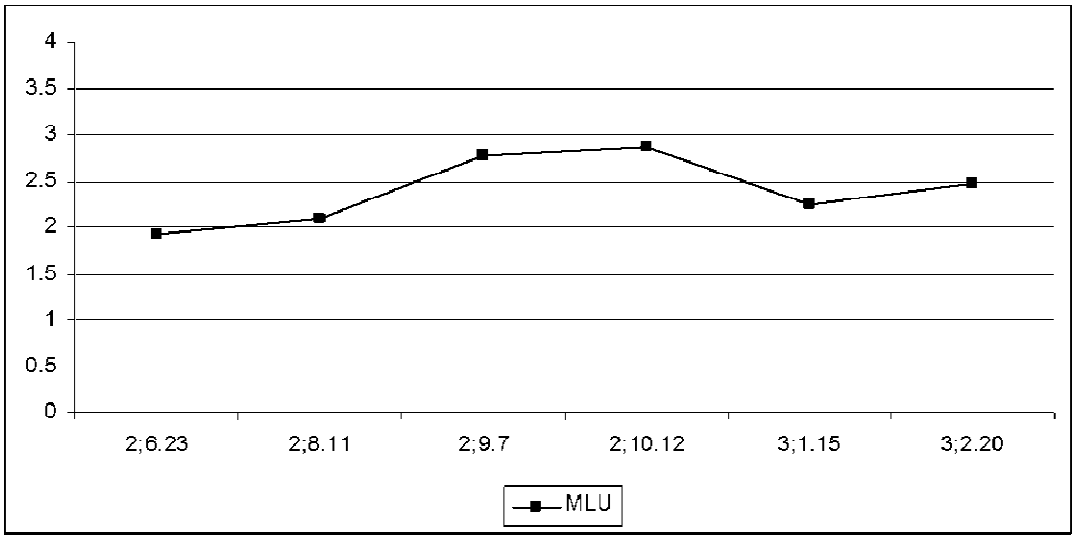
Age	U %	It %	En %	It-En %	En-It %	MLU
2;6.23	0	61.9	19	19	0	1.92
2;8.11	3.9	63.1	23.6	9.2	0	2.10
2;9.7	1.6	88.4	6.6	3.3	0.8	2.77
2;10.12	0	92.3	1.5	6.1	0	2.86
3;1.15	0	99	1	0	0	2.25
3;2.20	0	100	0	0	0	2.48

Figure 3.7 Overview of Matelda's data



The pattern of Matelda's language production shows that while initially almost 20% of her utterances are in English and almost 20% are mixed utterances, her use of English decreases considerably after age 3. It must also be taken into consideration that she spent a considerable amount of time (almost 10 weeks) in Italy just before she turned 3.

Figure 3.8 Matelda's MLU



Matelda's MLU grows from age 2;6 to 2;10, but it decreases after age 3.

3.7.2 Early linguistic development: one and two word utterances

At the time of her first recording, Matelda was starting to produce her first 2 word utterances. She frequently produced English or mixed utterances. The mixed utterances generally start in Italian. However, unlike the other children, she doesn't seem to use switching as a strategy to compensate for the lack of lexical knowledge. For most of the words she uses in English, she shows knowledge of the Italian equivalent.

3.53 *MAT: e poi this
and then this
*INV: che cosa?
what?
*MAT: è mine
it's mine M 2;6.23

3.54 *MAT: perche' vuole more carotine
because he wants more carrots M 2;9.7

3.55 *INV: George anche è un orso?
George also is a bear?
*INV: anche lui lo facciamo giallo
we also make him yellow
*MAT: o forse white, o forse bianco
or maybe white, or maybe white
*INV: l'altro conservalo per giocare con papá
keep the other to play with dad
*MAT: e dopo fai il book
and then do the book
*INV: devo farti il libro?
I have to do the book?
*MAT: yes
*INV: di pongo?
of clay
*MAT: yes

Matelda acquires inflection, gender and number agreement by 2;9, and soon after this stage we can observe a reduction in the rate of determiners omission. From age 3 she starts producing increasingly complex sentences. Even though her Italian is the dominant language in her environment, she seems to switch very often to English. Some of her utterances also show cross-linguistic influence. Sentence 3.56 is an example of the influence of English on the selection of the auxiliary in Italian. Matelda is looking at a picture of a bear who is annoyed

because a wasp is flying around him. The picture shows the bear moving his hands in the air to get rid of the wasp. The gesture is interpreted by Matelda as the hand movement that people do when they are too hot.

3.57 *INV: e perchè fa cosí?
and why does he do this?
*MAT: perchè è hot
because (it) is hot
*INV: perchè ha caldo
because he is hot

Matelda's use of English decreases after a long trip to Italy, after which she stops producing English sentences. At the stage of the last recording, Matelda is producing complex sentences that are the closest to adult-like sentences, often sounding like an imitation of adult's language.

3.58 *MAT: cosa facciamo di bello ancora?
what else are we going to do?

3.59 *INV: e col bianco che ci facciamo?
and what are we doing with the white?
*MAT: le corna
the horns
*INV: le zanne dell'elefante
the elephant's horns
*MAT: io le faccio
I'll do them
*MAT: poi si attaccano
then they are attached
*MAT: fai questo
do this

3.60 *MAT: questo non lo sai fare, vero?
you can't do this, can you? M 3;2.20

3.7.3 Comparison with other Italian data

Matelda's MLU and linguistic development resemble those of the Italian children from the CHILDES corpus and to Lina. Her MLU is higher than Lukas' and on average similar to Carlo's (initially her MLU is slightly lower than Carlo's, after age 2;9 it becomes slightly higher). On the basis of this comparison, she might be developing Italian as a strong language.

3.8 Chapter summary and conclusion

In this chapter I provided a general overview of the first phases of the development of Italian in four bilingual children, on the basis of spontaneous longitudinal data. The data collected for this study is compared to monolingual and bilingual data from other studies, based on two sets of children who acquired Italian as a minority language in Sweden and in the UK (section *3.3.7*). For each child, I have provided a brief introduction on the background and a summary of the linguistic use of each recording. The introductory section was followed by a more detailed analysis of the one and two word stage, with representative examples of some developmental phases. Finally, I used the results on the MLU and some observations on the linguistic development of Italian from Cipriani et al. (1993) Serratrice (1999) and Bernardini (2004) studies to make an initial prediction on the children's development of Italian as a weak or strong language.
The results show that there is quite a wide variation among the children, and in some cases their development seems slower, if compared to their monolingual or bilingual peers. Since this study does not follow the children after the age of 4, it is not possible to determine whether the structures that were not acquired up to that stage will never be acquired, or whether the process of acquisition is only temporarily slower. I would like to argue for the second position (also see Meisel 2007 and Bonnesen 2009), assuming that a bilingual child has the potential to acquire both languages given enough exposure to the input. The lack of use of the language may cause a slow rate of acquisition, which can finally result in the child's refusal to speak the language, or in his/her choice of English (or another majority language) for a more effective communication. As the analysis of the data demonstrates, all the children involved in this study use Italian when talking to their mothers, but some are faster at acquiring the language and can use it more productively. Some of the observations concerning particular aspects of the children's development constitute the basis for the analysis of the weak language, which will be discussed in the next chapter.

CHAPTER 4
Strong and weak development of Italian

4.1 Introduction

As I have shown in chapter 2, research has so far shown that the weak language is different in some aspects from the strong language, but there is not yet a full account on how to assess weakness at the early stages of linguistic development. The analyses of the weak language proposed so far have taken into account factors such as code-mixing, rate of acquisition of syntactic structures or functional categories, production of norm-deviant (or target-deviant) forms, vocabulary, lexical and verb types, avoidance of complex structures, MLU and Phonological MLU. The methodology I employ is based on the combination of the investigation of some factors that have been previously analysed and some that are new to the study of the weak language. First, I explain the methodology and the reasoning behind the choice of factors investigated, then I turn to the analysis of rate of acquisition, production of target-deviant forms, lexicon, code-switching, MLU and discourse pragmatics. Finally, I show how the results can be used to determine whether the children develop Italian as a weak or strong language by using the Weak Language Scale. Two important methodological choices are made in this analysis of the weak language: the use of spontaneous longitudinal data, which allows us to account for the development of Italian at different stages, and the use of monolingual and bilingual Italian data for comparison. This analysis of the weak language differs from previous ones in the factors analysed, the data used for comparison and also the creation of the Weak Language Scale. Since it is the first time that the weak language is analysed using this method, it is not yet possible to verify the universal validity of the results. However, the methodology employed can constitute the basis for further research.

4.2 *Weak vs. Weaker*

The use of the terminology in studies on language dominance is strictly dependent on the method of analysis. The terms that have been used so far in relation to the "different" or "less-developed" language are *weaker* or *non-dominant*. The two terms have been used interchangeably, but in my opinion the term *non-dominant* indicates the relationship between the two languages (one is dominant over the other), while the term *weaker* suggests the idea of "reduced ability" or "low proficiency" in comparison to the other more developed language.

As I have shown in chapter 2, researchers have analysed the weak language by drawing a comparison with the strong one, with the aim to determine whether there are differences in the development in the two languages in a bilingual child. My analysis is based on a different approach: since the aim of this thesis is to determine whether Italian develops as a strong or weak language, independently of the development of English, I use only Italian data from bilingual and monolingual children. Therefore, the purpose of this research is not to determine whether in the four bilingual children Italian is *weaker* or *stronger* than English, but if its development differs from that of other bilingual and monolingual Italian speakers. In order to stress the difference between the traditional approach that compares two languages in a single individual (such as French and German in the bilingual children studied by Bonnesen, 2009) and my own methodology, I use the terms *weak* and *strong*.

As Cantone et al. (2008) suggest, it is essential to use as much data as possible to set a monolingual and a bilingual norm (see chapter 2, section 2.4). There are several studies that present large amounts of data from monolingual Italian children, but not much data is available from bilingual children acquiring Italian in minority contexts. In the analysis that follows I compare – when available and appropriate – the results of the four bilingual children to those of monolingual and bilingual children from previous studies. The analysis presented in this chapter can be considered a starting point for a possible future large-scale study of the development of Italian in minority contexts and for the creation of a "bilingual norm" for Italian in different contexts of acquisition.

To summarise, the main difference between this study and the previous ones lies in the type of data and in the selection of criteria employed to determine weakness. Previous studies have compared two different languages, while I compare data from

different groups of children focusing only on Italian. Another difference lies in the use of the terminology. Since the terms *weaker*, *stronger*, *dominant* and *non-dominant* imply a relationship between two languages, I am going to use the terms *weak* and *strong* when referring to the development of Italian. The use of *weak* instead of *weaker* is aimed at stressing the difference in the methodological approach.

4.3 Methodology

As I have shown in chapter 2, several factors have been proposed in the assessment of the weaker language, affecting the types of analysis presented. In many studies on bilingual development there is some reference to dominance or to the child's proficiency in each language, but there is not yet a unified definition of *weak language* that could be applied cross-linguistically. In this chapter, I propose a method of analysis that applies to Italian, but that can be adapted to other languages. In order to have a broader overview of the development of Italian, I employ both the data collected for this thesis and data from previous studies.

The first important methodological choice concerns the selection of factors to be taken into account to determine whether a language is weak or strong. On the basis of previous studies (see chapter 2), I selected some of the factors that I considered relevant and adapted the analysis to the Italian language. In addition, I introduced the analysis of discourse pragmatics, which has not been previously tested in weak language studies. The factors selected are *rate of acquisition*, *code-mixing*, *vocabulary*, *use of target-deviant forms* and *discourse pragmatics*. These factors have been chosen because they cover different areas of linguistic competence, they can show a comprehensive and longitudinal view of the child's linguistic development and they have been widely studied in the literature on monolingual and bilingual language acquisition.

The second important methodological choice concerns the selection of comparative data. In this chapter, I mostly use the results from the analyses carried out by Serratrice (1999), Bernardini (2004) and Cipriani et al.(1993), in order to present a wider perspective on the development of Italian[30]. As I have shown in chapter 3, the Italian data from these three longitudinal studies can be used for comparison

[30] Results from other studies are also used (see section 4.4.1).

purposes, since it covers an age range similar to that of the children analysed in this thesis and also because both Serratrice's and Bernardini's bilingual studies give an indication of the children's linguistic dominance (see chapter 3).
In the sections that follow, I present the analysis of each factor (section 4.4-4.9) and then I combine the results introducing the Weak Language Scale (section 4.10-4.11).

4.4 Rate of acquisition

The majority of studies on the weak language have found that this language develops slower than the strong language (Schlyter 1993, Meisel 2007). If we assume that the strong language is like an L1, we can conclude than the weak language may develop slower than an L1, and it is therefore possible to use monolingual data to compare the results. To further verify the results from this comparison, it is also possible to analyse other bilingual data.
Several studies have shown that all children follow similar patterns of development and reach specific milestones approximately at the same age (Guasti 2004). Research on linguistic dominance has shown that even though the weak language follows the same developmental milestones as the L1, it can develop at a slower rate. This phenomenon has been observed mostly by comparing MLU values or the acquisition of specific syntactic properties (Bonnesen 2009). In order to analyse the rate of acquisition, we should take into account at least one factor on the basis of which children are comparable, such as age, MLU, developmental phases (Guasti et al. 2003) or parameter setting, depending on the variable we consider the most relevant for the specific purpose. If a child's production falls outside the average time-frame, we may conclude that there is a slow rate of acquisition, at least in the domain we are analysing.
From the analysis of the longitudinal data (see chapter 2), it emerges that overall there is some degree of variation in rate of acquisition among the four children. For example, Paolo is the child who starts producing Italian utterances the latest and with an irregular developmental trend in comparison to the other children.
In order to determine the rate of development, I will analyse the acquisition of articles for two main reasons. Firstly, articles have been widely studied both in the monolingual and bilingual acquisition of Italian (Bottari et al. 2001, Guasti et al. 2003, Bernardini 2004, Kupisch and Bernardini 2007, Kupisch 2007, Ferrari and Matteini 2008). Secondly, they are very frequent in the input and in Italian they are

often obligatory, and this provides us with many opportunities to study their acquisition. Articles are also the most recurrent forms in the Italian language (as well as other languages) and they are often obligatory. Another reason for choosing to analyse the acquisition of articles is that their omission is a typical developmental phenomenon in Italian as well as other languages (Guasti et al. 2003). The availability of data for comparison, the frequency in the language, and the universality of the phenomenon of omission make articles suitable for examining the rate of acquisition.

4.4.1 The acquisition of articles in Italian

In Italian, definite and indefinite articles can precede nouns (il tempo-*the time, the weather*), adjectives (il bel tempo-*the nice weather*) and verbs (il passare del tempo-*the passing of time*) and they agree in gender and number with the word that follows. The choice of article depends on the phonetic component of the word that follows, as I show in tables 4.1 and 4.2.

Table 4.1 Definite articles in Italian

	Singular/Plural	Examples
Masculine	il/i lo/gli (words starting with *s*+consonant, z, ps, gn, x, y) l'/gli (words starting with a vowel)	il libro/i libri lo studente/gli studenti l'uomo/gli uomini
Feminine	la/le l'/le (words starting with a vowel)	la classe/le classi l'idea/le idee

Table 4.2 Indefinite articles in Italian

	Singular/Plural	Examples
Masculine	un/dei uno/degli (words starting with *s*+consonant, z, ps, gn, x, y) un/degli (words starting with a vowel)	un libro/dei libri uno studente/degli studenti un uomo/degli uomini
Feminine	una/delle un'/delle (words starting with a vowel)	una classe/delle classi un'idea/delle idee

According to Cipriani et al. (1993), the acquisition of determiners[31] (including articles) can be divided into 4 stages. The first stage is characterised by a prevalence of omitted forms, the second by a decrease in omission rates and increase in errors[32], the third by a significant decrease in omission rates and the fourth by a high percentage of correct use of articles, and a few errors with plural articles. Their data show that articles emerge between age 1;8 and 2;2 and omission decreases considerably between age 1;8 and 2;7. Other data from monolingual Italian children show that determiners can emerge at the age of 1;8, but they are used productively in obligatory contexts only after age 2;2 (Antelmi 1997: 78). Antelmi's analysis shows that errors such as gender and number marking are very infrequent in determiners, and they completely disappear after age 2;5. However, the percentage of omission is higher than the production of correct forms until the age of 2;1. This proportion shifts after age 2;3, when an increasing number of determiners is produced in obligatory contexts.

In a study on the use of determiners across languages, Guasti et al. (2003) analyse article omission not on the basis of age or MLU, but of "periods of linguistic development" (p. 4), that correspond to three stages of lexical acquisition (1-100 words, 101-200 words, and more than 200 words). Their research shows that Italian children start producing their first determiners at age 2, and the percentage of omission drops from 52% at the first stage to 17% at the second stage[33]. Guasti (2007) also observes that children produce few determiners in correspondence to an MLU value between 1 and 1.5, and the rate increases when they reach an MLU between the values of 1.5 and 2.5.

The studies of bilingual children show quite a similar pattern. Carlo starts producing definite articles at age 1;10 (MLU 1.165) and he progressively produces more articles at 2;0.1 (MLU 1.178), with a considerable increase between age 2;5.26 (MLU 2.631) and 3;0.17 (MLU 3.306) (Serratrice 1999). His data show high omission rates until the age of 2;2.17 (MLU 2.007), after which they progressively decrease. Bernardini (2004) shows that Lina's omission of determiners starts decreasing at age 2;8.21, and Lukas' at age 2;11 (see table 4.3).

[31] See also chapter 2 section 2.3.

[32] What they consider as error is the wrong agreement between determiner and noun (*un lucertola* – masculine article and feminine noun) or the use of two determiners, generally an indefinite followed by a definite article (*una l'aquila*).

[33] Guasti et al.'s research is based on the data from Raffaello, Martina and Diana from the CHILDES corpus.

Table 4.3 shows the age of production of first determiners and the age of decrease of omission for each child. On the basis of Cipriani et al. (1993), I consider 'decrease of omission' the stage in which at least 75% of determiners appear in obligatory contexts.

The table shows that there is not a significant difference among the children in terms of age of first production. However, the age of decrease of omission oscillates between 1;8 and 2;11.

Table 4.3 Articles: age of first production and age of decrease of omission[34].

Study	Child	Age of production of first determiners	Age of decrease of omission of determiners
Antelmi (1997)	Camilla – IT	1;10	2;0
Cipriani et al. (1993)	Rosa - IT	2;0	2;8
	Martina - IT	1;8	1;11
	Diana - IT	1;8	1;8
	Viola - IT	2;0	2;5
	Guglielmo - IT	2;2	2;4
	Raffaello – IT	2;0	2;4
Ferrari and Matteini (2008)	Sabrina – IT	1;11 (or earlier)	2;3
Bernardini (2004)	Lukas – IT-SW	2;2	2;11
	Lina – IT-SW	2;2	2;8
Serratrice (1999)	Carlo – IT-EN	1;10	2;2

4.4.2 The acquisition of articles in the 4 bilingual children

I will proceed by looking at the production of definite and indefinite articles in the data from the 4 bilingual children. Following Guasti et al. (2003), in this analysis I include only articles in sentences containing a verb. This method greatly reduces the

[34] In the studies reported in table 4.1, definite and indefinite articles are included in the analysis.

number of utterances considered, especially at the early stage, but it is a more effective way to determine whether there is an actual omission of determiners[35].
The following tables show the number of articles produced in obligatory contexts, the number of obligatory contexts, and the percentage of omitted articles.

Table 4.4 Articles - Matelda

Age	MLU	Number of articles produced in obligatory contexts	Percentage of omitted articles
2;6.23	1.92	0/0	0%
2;8.11	2.1	2/2	0%
2;9.07	2.77	11/19	42.1%
2;10.12	2.86	9/10	10%
3;1.15	2.25	2/2	0%
3;2.20	2.48	5/5	0%

Table 4.5 Articles – Paolo

Age	MLU	Number of articles produced in obligatory contexts	Percentage of omitted articles
3;1.27	2.11	1/1	100%
3;3.23	1.81	5/9	44.44%
3;4.25	1.77	0/5	100%
3;7.10	1.66	0/3	100%
3;10.19	3.02	12/15	20%
3;11.17	2.65	11/17	35.29%
4;0.29	2.87	10/17	41.17%
4;1.28	2.21	11/15	26.66%

[35] The study carried out by Guasti et al. (2003) aimed at establishing the factors influencing the omission of determiners comparing data from Italian, Catalan and Dutch.

Table 4.6 Articles – Costanza

Age	MLU	Number of articles produced in obligatory contexts	Percentage of omitted articles
1;1.16	1.78	0/0	0%
1;12.10	1.61	0/0	0%
2;2.17	2.87	7/7	0%
2;4.09	2.98	18/19	5.26%
2;6.07	2.58	10/10	0%
2;7.16	2.86	13/13	0%
2;9.14	4.24	32/32	0%

Table 4.7 Articles - Francesca

Age	MLU	Number of articles produced in obligatory contexts	Percentage of omitted articles
2;4.20	1.39	0/0	0%
2;5.10	1.26	0/0	0%
2;6.19	2.06	2/5	60%
2;7.28	2.21	3/4	25%
2;9.07	2.11	6/19	68.42%
2;10.17	2.21	3/14	78.57%
3.0.17	2.45	3/9	66.66%
3;1.17	2.80	24/33	27.27%
3;2.27	3.13	22/31	29.03%
3;5.0	2.84	13/17	23.52%

These results show that there is variation among the children both if we compare them on the basis of MLU and on the basis of age. This variation is also shown in the analysis of the age of first production and that of decrease of omission[36] (Table 4.8). If we compare the results in table 4.8 with those of table 4.3, we can see that there are more differences between monolinguals and bilinguals than among monolinguals.

Table 4.8 Determiners: age of first production and age of decrease of omission

Child	Age of production of first determiners	Age of decrease of omission of determiners (75%)
Francesca	2;7.28	3;5.0
Paolo	3;3 (or earlier[37])	4;1.28 (74%)
Matelda	2;9.07	2;10.12
Costanza	2;2.17	2;2.17

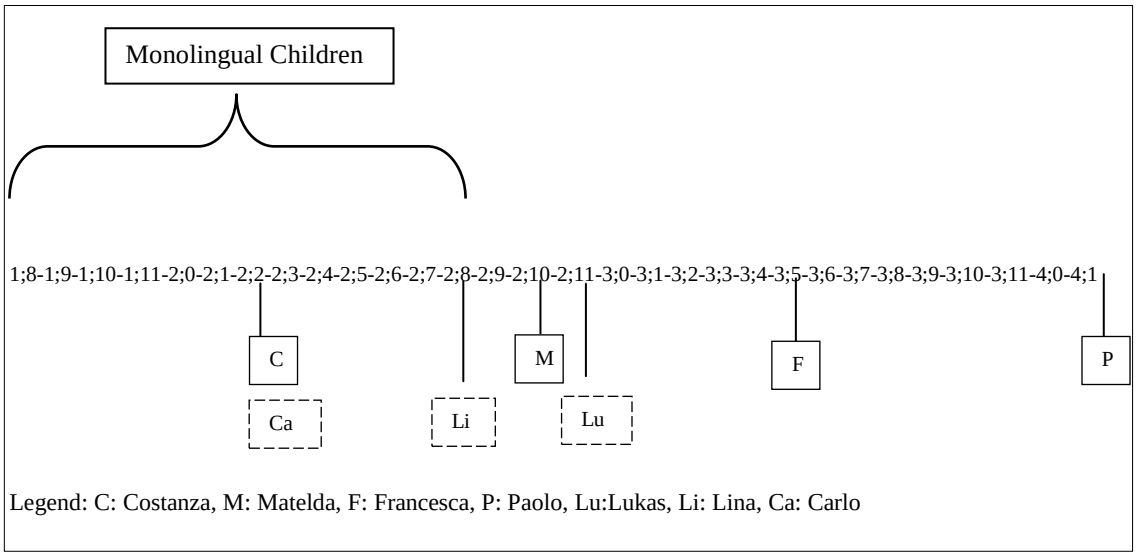
The average time between first production and decrease of omission for monolinguals is about 3 months. The data from Costanza and Matelda show similar values to monolinguals, while Paolo and Francesca acquire determiners over 10 months.

Figure 4.3 shows the difference among the children, focusing on the age of decrease of omission. The results contained in the curled bracket are those of the monolingual children and the ones below are those of the bilingual children (see legend).

[36] Paolo does not reach 75%, but his closest result is 74% at the age of 4;1.

[37] Paolo already produces determiners in the first recording.

Figure 4.1 Ages of decrease of omission of articles across different groups



This figure clearly shows the difference between the stage at which each child stops omitting articles. If we observe the bilingual data, we can find significant differences in the stage of omission. This finding confirms that the rate of acquisition is a factor that can help us differentiate the children's development and it can be considered significant in the study of the weak language. Another important finding is that three of the bilingual children (Carlo, Costanza and Lina) perform similarly to monolingual children. It is also interesting to note that according to Bernardini and Serratrice, Lina and Carlo develop Italian as a strong language.

For the purpose of the final analysis, the acquisition of articles will be regarded as the percentage of presence of articles in obligatory contexts. As I will show in section 4.10, this method is used in order to account for the rate of acquisition at the different stages under examination. The fact that four bilingual children perform differently from monolinguals might be a first sign of their weak development. However, it is necessary to examine other factors to have a more comprehensive view of the children's development.

Slow rate of acquisition has been analysed here on the basis of the acquisition of determiners, but it could be extended to other categories and to parameters that are set early and uniformly.

4.5 Production of target-deviant forms

Previous studies of the weak language have associated the high production of target-deviant forms (or errors) to weak language development (Schlyter 1993, Döpke 2000, Bonnesen 2009). Research on monolingual acquisition has shown that children make few errors when they acquire the L1[38]. Most of these are considered to be developmental errors, which are common to all children acquiring that specific language. An example of developmental error in Italian is the omission of articles, which I have discussed in the previous section. All target-deviant forms can be considered useful to determine whether a child develops Italian as a weak or strong language, because, as the previous analysis shows, errors may continue to be produced even after a stage in which they are expected to disappear. In this analysis I consider target-deviant forms affecting word-order, gender and number agreement and verb inflection, without making a judgement as to which of those constitute a developmental error. The prediction is that the weak language will exhibit target-deviant forms at all stages of development, and some of these forms might persist in the child's production. The repeated presence or the fossilisation of the errors can be attributed to the properties of the input or to the lack of use of the language: while developmental errors in monolingual children are replaced by the correct form through exposure to positive input (see section 2.2), bilingual children who acquire a language in a minority context might not have access to a sufficient amount of input required to acquire the correct form as fast and efficiently as monolinguals.

The data reported in Table 4.9 show the number of target-deviant forms in each recording. Mixed and unintelligible utterances are not considered (see Appendix E for more details). Since the number of recordings is different for each child, I include the average number of target-deviant forms per recording.

[38] See O'Grady (1997) and Guasti (2004) for a discussion on developmental errors.

Table 4.9 Number of target-deviant forms

Child	word order	gender agreement	number agreement	verb morphology	total	Average per recording
Francesca	5	13	3	35	56	5.6
Costanza	1	1	0	6	8	1.14
Paolo	9	12	0	18	39	4.87
Matelda	0	1	2	6	9	1.28

As shown in Appendix E, the number of errors increases along with the MLU only in Francesca's and Paolo's data. Even though the number of target-deviant forms is low compared to the number of correctly formulated utterances, the results shed further light on the differences among the children. The number of target-deviant forms is also inversely proportional to the overall input the children receive, as I will show in chapter 6. This correspondence might suggest that the children who are exposed to a more limited input might not receive the necessary positive evidence to develop a strong language. It has been demonstrated that children acquire the grammatical rules of their language not by being corrected, but by being exposed to a sufficient amount of input showing the correct use of language. In bilingual contexts, it is possible that the children who do not receive sufficient input in the minority language produce target-deviant forms over a longer period than monolinguals, and these forms might get temporarily fossilised, as shown in the examples presented in chapter 3 (section *3.4.2*).

The analysis of errors, as the one on the rate of acquisition, proves to be significant in determining differences among the bilingual children and it will therefore be included in the final analysis of the weak language. Serratrice's and Bernardini's studies do not provide data on the production of all the types of errors that I examined, therefore they cannot be used for comparing the results.

4.6 Vocabulary

Another factor that has been widely analysed to study the weak language is the acquisition of the lexicon. The lexicon is a domain that undergoes constant expansion. Children generally produce their first 50 words at about 18 months; between 18 and 24 months their vocabulary rapidly increases and they acquire up to

9 words a day, while they start to associate words with their meaning increasingly fast (Guasti 2007: 125).

The first reason for analysing the lexicon in this study is to determine the differences among the children in the growth of the vocabulary. The second and more important reason for the purpose of this thesis is that, while the grammar is an inborn structure in the mind, the vocabulary has to be acquired from the external environment, therefore the input plays a major role in lexical acquisition.

Several studies have analysed the acquisition of the lexicon in Italian children (Caselli, Pasqualetti and Stefanini 2007, D'Odorico et al. 2001). D'Odorico et al (2001) show that the early acquisition of the lexicon can be subject to quite a degree of variation: at 19 months about 40% of children produce about 50 words, and only 30% of children produce more than 100 words. They also found that the most significant factor in the rate of lexical development is the mother's education, further showing that the input has a major role in lexical development. Most children produce at least 200 words at age 2, but other produce only 50 words at this age.

Bilingual children living in a predominantly monolingual environment may have a more limited lexicon in the minority language. Moreover, children who are mostly exposed to the minority language at home and with only one of their parents, are likely to develop lexical areas specific to their needs and to the activities they perform with the adult.

In this section, I show the children's vocabulary size in each recording, considering the number of word roots produced for each lexical category - nouns, adjectives, verbs - (see Appendix F for the complete list of lexical items). It has to be taken into consideration that this type of analysis does not reflect the actual full lexical knowledge of the child, since it is based on a small sample of the child's lexical production, which reflects only part of the child's knowledge[39].

[39] An addition to the longitudinal data could be the employment of a lexical test. "Il primo vocabolario del bambino" is the Italian version of the "MacArthur-Bates Communicative Development Inventories – CDI" (Caselli and Casadio 1995, Caselli, Pasqualetti and Stefanini 2007). This questionnaire has been administered to more than 700 monolingual children from the earliest developmental stage up to 36 months. Caselli, Pasqualetti and Stefanini (2007) argue that the questionnaire, filled out by the parents, is a reliable source of information, since parents have been shown to reflect quite accurately their child's performance. However, it is possible that the ability of reporting the bilingual child's performance can be reduced in the bilingual parent. Moreover, to show the actual development, it would be necessary to keep a diary and fill out the test at least once every month. Given these premises, I am aware of the limitations of the analysis based only on the data from the recordings. The analysis of the lexical production of each child takes into account the number of nouns, adjectives and verbs.

Figures 4.2-4.5 show the number of different word roots in each recording (immediate repetitions and incomplete words are excluded).

Figure 4.2 Vocabulary: Matelda.

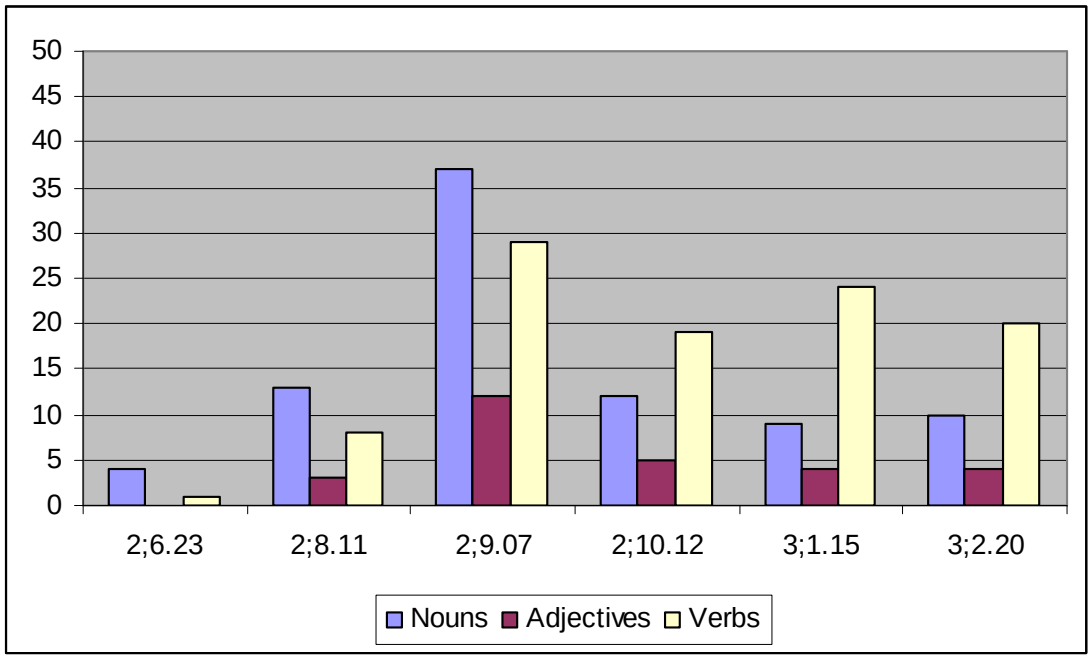


Figure 4.3 Vocabulary: Paolo

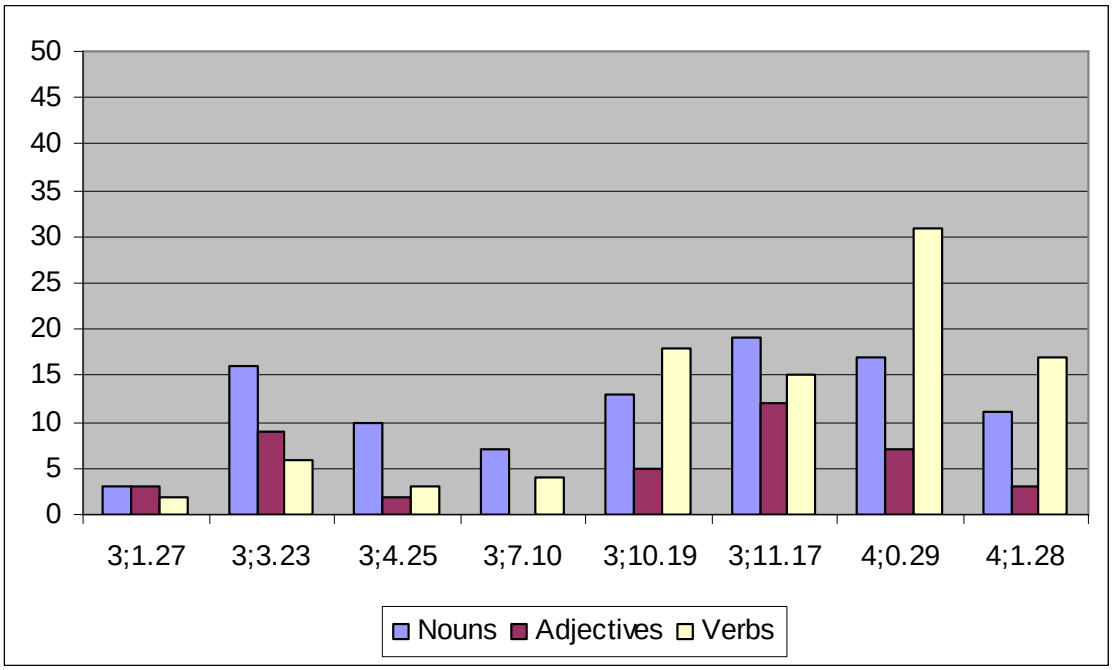


Figure 4.4 Vocabulary: Francesca

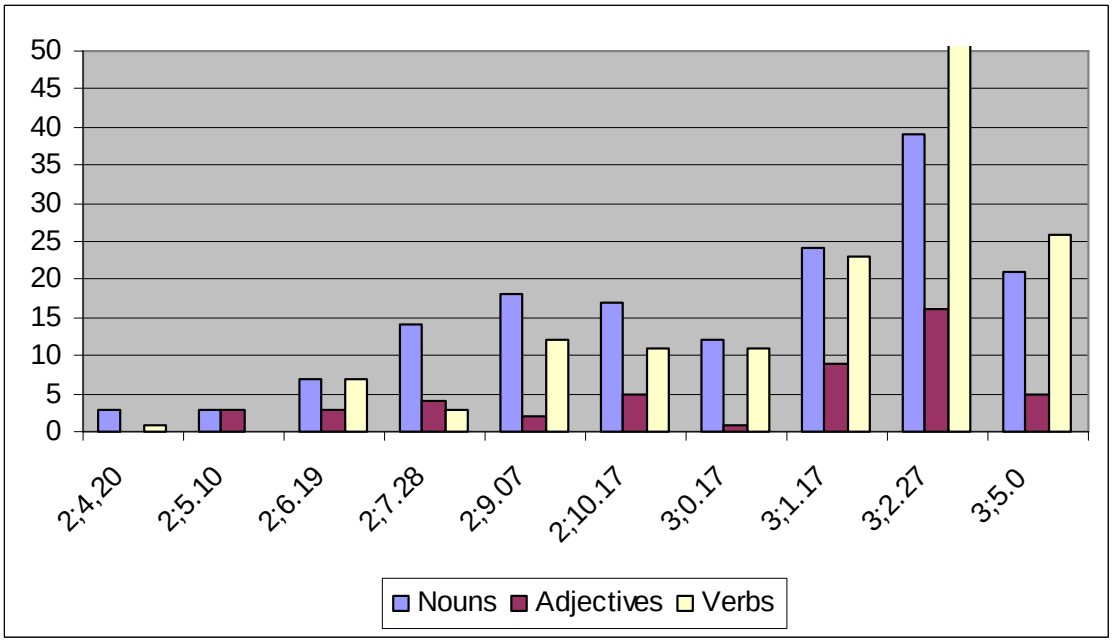
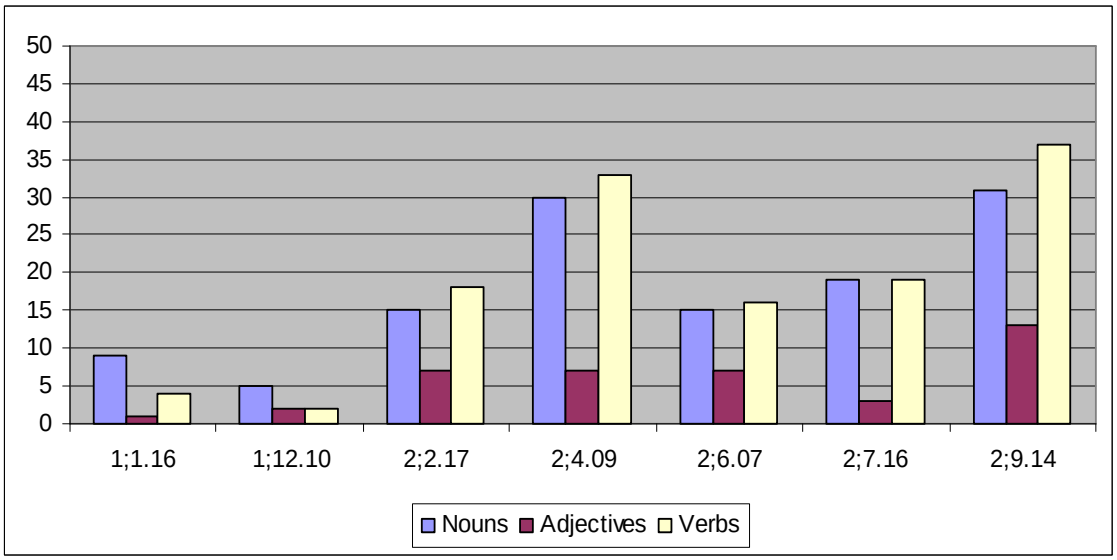
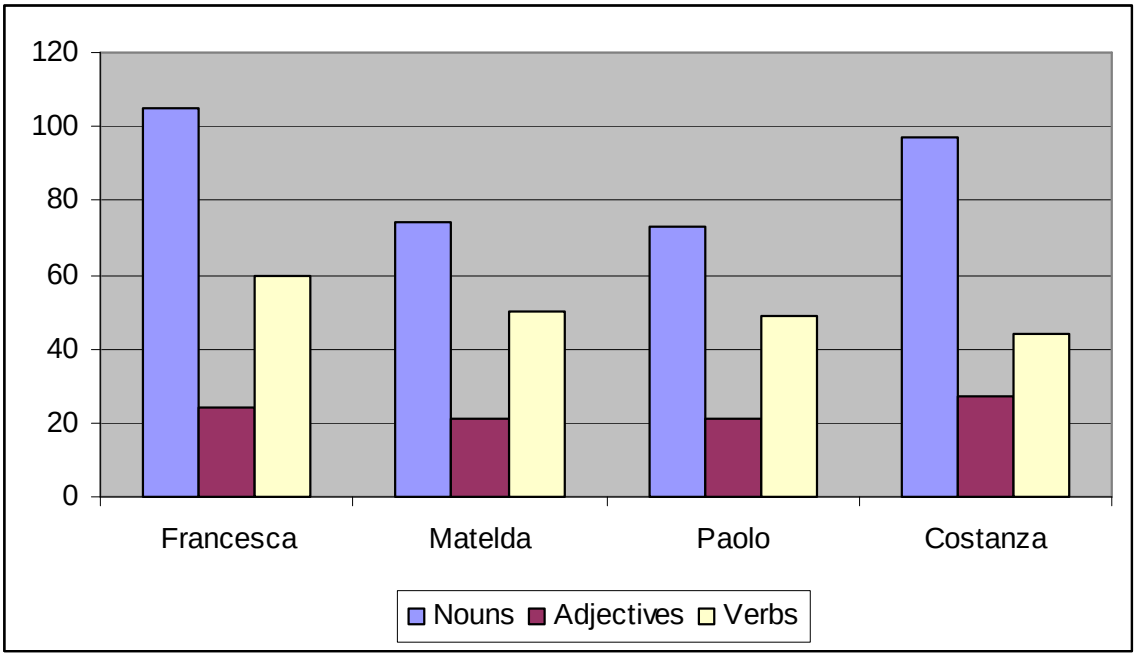


Figure 4.5 Vocabulary: Costanza



This first analysis takes into account the amount of lexicon in each recording in order to show the developmental pattern of the lexicon. The second analysis (Figure 4.6) shows for each child the total amount of lexical items. In this analysis, each word root is calculated only once.

Figure 4.6 Total vocabulary



It has to be considered that the total vocabulary has been calculated on the basis of the recordings available, and that not all children were recorded with an equal frequency. I have therefore divided the number of total different words for the number of recordings, to show the average vocabulary size on the basis of the number of recordings (Table 4.10).

Table 4.10 Average vocabulary size in each recording

Francesca	Matelda	Paolo	Costanza
18.9	24.16	17.87	24

These results again show variation among the children, and resemble the results obtained in the analyses of the rate of acquisition and the production of target-deviant forms.

In order to compare these results to other monolingual and bilingual studies, it is necessary that they also employ spontaneous data. While bilingual studies of Italian do not provide large-scale results on the lexical development based on spontaneous data (to my knowledge), there is normative data on monolingual Italian children based on standardised tests (Caselli and Casadio 1995, Caselli, Pasqualetti and Stefanini 2007). Due to the nature of these tests, it is not possible to compare the results to those obtained by analysing spontaneous data. However, the two

methodologies (collection of spontaneous data and lexical test) can be combined to further assess the child's lexical production.

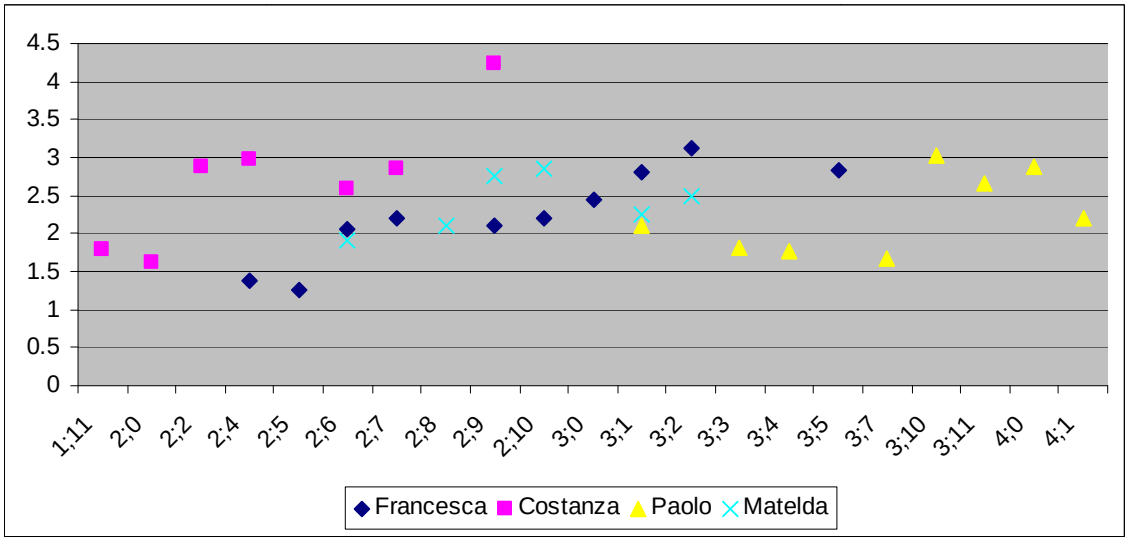
4.7 Mean Length of Utterance

The analysis of the Mean Length of Utterance (MLU) has been used since the 1970s (Brown 1973) in developmental linguistics research. The calculation of the MLU is used to determine the average number of words or morphemes that a child uses in each utterance produced. As the child's linguistic ability develops, the vocabulary grows, the syntactic structures used become more complex and children produce increasingly long sentences. Therefore, the calculation of the MLU, together with other factors, can be employed to provide further evidence on the rate of a child's linguistic development.

The average MLU is calculated in the same way across languages. However, some languages are more morphologically complex than others, and their complexity results in longer average MLU values, especially if the MLU is calculated in morphemes. For this reason, when comparing two languages, researchers often calculate the MLU in number of words rather than morphemes, and this is the methodology adopted here.

Figure 4.7 and 4.8 show the MLU values and average MLU of all children. The values are calculated summing the total number of words in each sentence and dividing the result by the number of utterances. The reason for calculating the MLU in words is the possibility to compare the data to the other studies that so far I have used for comparison. Bernardini and Serratrice analyse the MLU in words rather than in morphemes in order to show the difference between Italian and the other language (English and Swedish).

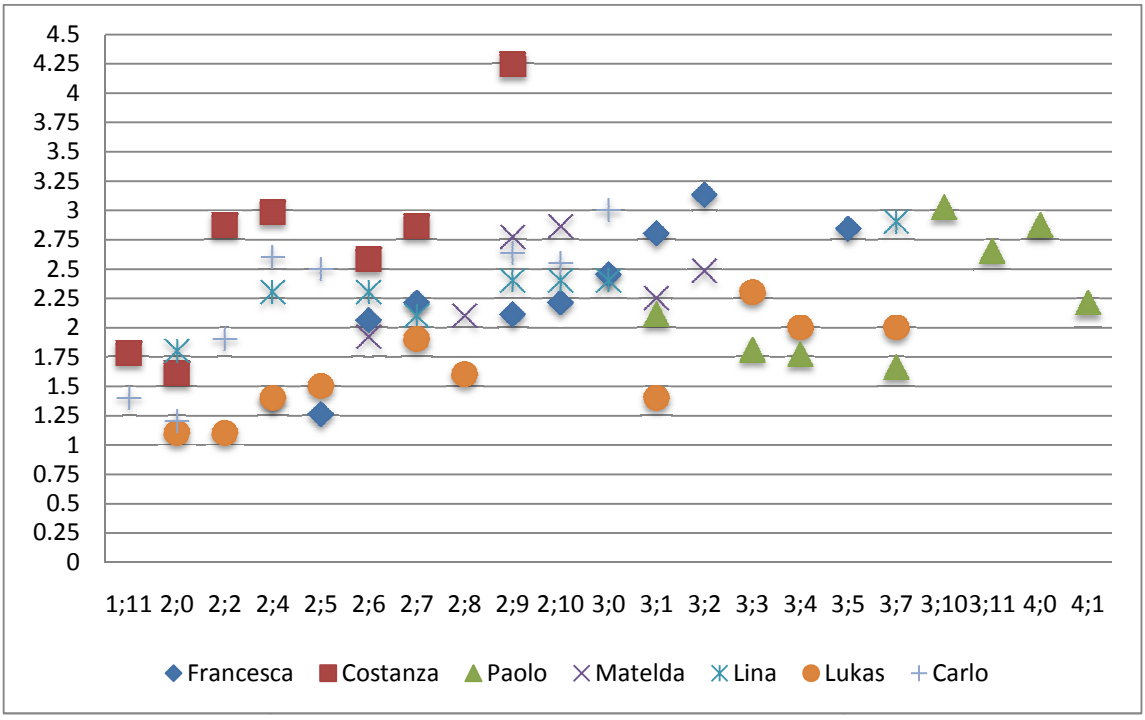
Figure 4.7 Children's MLU



As Figure 4.7 shows, while Costanza's, Francesca's and Matelda's MLU increases almost constantly, Paolo's MLU seems to decrease from age 3;1 to 3;8 and reach the highest peak after age 3;10, after which it decreases.

Figure 4.8 shows the data from the four bilingual children along with that of Lina, Lukas and Carlo (the children from Serratrice's and Bernardini's studies).

Figure 4.8 Comparison of average MLU of bilingual children



The figure shows that Lukas, the child who develops Italian as a weak language, has the lowest average MLU. We can observe that at three stages (age 3;3, 3;4 and 3;7) Paolo's average MLU is even lower than Lukas' and in two earlier stages it is lower than Francesca's and Matelda's. This first observation can lead us to assume that if the MLU is a valuable measure to assess weak language development, Paolo and Lukas are developing Italian as a weak language. On the other hand, the analysis presented so far has shown that Costanza, Carlo and Matelda develop Italian similiarly to monolingual children. In the figure above, we can observe that Costanza reaches the highest value of MLU from the earliest stages, while Carlo, Lina, Matelda and Francesca achieve similar results through the various stages.

The analysis of the monolingual data from Cipriani et al. (1993) shows that between 2;0 and 2;3 children have an average MLU ranging between 2.1 to 2.6 (with the exception of Diana, who reaches 4.1), between 2;4 and 2;7 the average MLU ranges between 1.9 and 3 and between 2;8 and 3 it ranges between 2.9 and 4.1[40].

From the comparison of the monolingual and bilingual data, it emerges that Lukas and Paolo have a lower average MLU than monolinguals and also bilinguals, while Costanza and Matelda achieve the results comparable to monolinguals. Francesca's MLU is initially low, but it constantly increases and finally reaches a value that is comparable to that of her monolingual and bilingual peers. Like the previous factors examined, the MLU is a valuable measure to compare children at different stages of development. In the analysis of the weak language presented in this chapter, the MLU will be used to establish four phases of development on the basis of which the children can be grouped and each factor can be analysed longitudinally.

4.8 Code-switching

Some researchers have regarded code-switching and code-mixing as a phenomenon connected to language dominance. The main argument is that the directionality of mixing can predict dominance, since children are more likely to mix from the dominant to the non-dominant language more often. Genesee et al. (1995) found that children with an unbalanced development seem to mix more when using the non-

[40] The results from monolingual children are not included in the figure to avoid confusion.

dominant language. Other studies, however, have presented evidence showing that mixing is not necessarily related to dominance or to its directionality. (Cantone 2007, Cantone et al. 2008, Müller 2008). In my analysis, I consider both the directionality of mixing and the percentage of mixed utterances over the total utterances produced.

Table 4.11 Code-switching

F It-En %	**F En-It %**	**C It-En %**	**C En-It %**	**P It-En %**	**P En-It %**	**M It-En %**	**M En-It %**
5.5	0	2	0	2.7	6.1	19	0
4.8	8	0	0	2	4.7	9.2	0
2.1	6.5	1.3	0	6	9	3.3	0.8
1.5	0	0	0.6	4	4	6.1	0
1.4	0.7	1	0	0	0	0	0
1.8	0.9	0	0	0	0.5	0	0
3.5	0	0.6	0	0	0	-	-
0.8	0	-	-	0	0	-	-
0.6	0	-	-	-	-	-	-
0	0	-	-	-	-	-	-
Average 2.2%	average 1.6%	average 0.7%	average 0.08%	average 1.8%	average 3%	average 6.2%	average 0.1%

The results on the directionality of mixing do not show sufficient evidence to consider this factor relevant in the analysis of the weak language. The highest percentage of mixing can be found in the data from Matelda, who switches in the majority of cases from Italian to English. On the basis of these results, I analysed the parents' data to see if there was a relationship between mixing and input and I found that Matelda's mother sometimes produces mixed utterances by ending an Italian sentence in English. This means that Matelda probably uses mixing more often than the other children because she replicates her mother's linguistic behaviour. I decided not to include code-mixing in the analysis of the weak language firstly because I believe that it is a communicative strategy that cannot be directly associated to weakness. Moreover, code-switching is only used by bilinguals, therefore in this case the comparison with monolinguals would not be possible.

4.9 Discourse pragmatics

Discourse pragmatics has not yet been analysed in studies on the weak language. While the previous factors examined in this chapter reflect the children's development of the lexicon and syntax, the analysis of the acquisition of subjects tests the child's discourse-pragmatics competence.

Since the appearance of the principles and parameters theory, the syntax of subjects has been widely investigated both from a theoretical and a developmental perspective. Many studies aimed at explaining the syntactic principles governing null-subject languages (Hyams 1986, Jaeggli and Safir 1989, Rizzi 1982, 1994, Holmberg and Roberts 2009). As I have shown in chapter 2, more recently, several studies have analysed the distribution of subjects from a developmental perspective (Serratrice 2005, Belletti, Bennati and Sorace 2007, Sorace et al. 2009).

Italian is a null subject language with a canonical SVO word order (see chapter 1, section 1.3). In general, null subjects are used where the verbal morphology disambiguates the reference (4.1).

4.1 Ho passato l'esame.
(I) passed the exam.

The use of an overt subject in a sentence such as 4.1 would either result in redundancy (4.2) or have a contrastive function[41] (4.3). Generally, overt preverbal subjects are used in contexts where they are marked as topic.

4.2 *Io ho passato l'esame.
I have passed the exam.

4.3 Io ho passato l'esame, Gianni no.
I passed the exam, Gianni didn't.

In terms of sentence felicity, when the subject is null, first and second person referents are hardly ever ambiguous, while third person referents can be ambiguous if the referent is not easily identifiable. This is for example the case of subordinate clauses containing a verb that could be referring to more than one

[41] These sentences have to be read with a neutral stress.

antecedent. The ambiguity is resolved by using a pronoun or by providing more information on the referent. It is also possible to disambiguate the reference by using gender (4.7) or number marking (Serratrice 2005).

4.4 Gianni$_1$ mi ha detto che *pro*$_1$ non ha passato l'esame.
Gianni told me that (he) didn't pass the exam.

4.5 *Gianni$_1$ mi ha detto che lui$_1$ non ha passato l'esame.
Gianni told me that he didn't pass the exam.

4.6 Gianni$_1$ mi ha detto che lui$_1$ non ha passato l'esame e Maria sí.
Gianni told me that he didn't pass the exam and Maria did.

4.7 Paolo$_2$ e Maria hanno fatto l'esame. Gianni$_1$ ha detto che lui$_2$ non l'ha passato.
Paolo and Maria did the exam. Gianni said that he hasn't passed it.

Sentence 4.5 is ambiguous, since the pronoun is redundant if Gianni is the referent of both verbs. The sentence would be felicitous if *lui* referred to another antecedent or if the subject represented focus (4.6). These examples show that the presence of overt subjects is not optional, but it responds to specific syntactic and discourse constraints. As Serratrice (2005) explains, *identifiability* and *accessibility* are two central concepts in discourse pragmatics. These two notions refer to the knowledge shared by the speakers and the ability to recover information in the hearer's memory. In Italian, first and second person pronouns are more likely to be null because the referent is identified by the contexts and they are used to express new information or focus. Third person subjects can be expressed not only by pronouns, but also by proper nouns, other NPs, and demonstratives and they can be ambiguous if they do not meet the criteria of identifiably and accessibility required in the discourse.

It has been demonstrated that the selection of null or overt subjects requires the activation of syntactic and pragmatic knowledge, and bilingual children who are acquiring a pro-drop and a non-pro-drop language may exhibit difficulties at this interface level:

> If the syntax–pragmatics coordination task is demanding for monolingual children, it can be twice as daunting in the case of bilingual children who have to map a larger array of language-specific morphosyntactic constructions onto a restricted set of language-universal pragmatic principles.
>
> (Serratrice, Sorace and Paoli 2004: 184)

As this and other studies show, this difficulty may result in cross-linguistic influence at the interface between syntax and discourse-pragmatics (Hulk and Müller, 2000; Müller and Hulk, 2001). The comparison between Carlo and the monolingual children in Serratrice, Sorace and Paoli's study shows that Carlo omits subjects less than his monolingual peers and, according to the authors, he produces overt subjects where null subjects are required.

On the basis of the findings from the studies mentioned above, I include discourse-pragmatics efficiency as a factor to be considered in the analysis of the weak language, assuming that the children who develop Italian as a strong language make a higher number of pragmatically correct choices in the use of subjects.

In the analysis that follows, I show the total number of inflected verbs produced in each recording and the number of correct and incorrect pragmatic choice of subjects.

Table 4.12 Use of subjects- Matelda

Age	Total inflected verbs	Correct pragmatic use	Incorrect pragmatic use
2;6.23	2	1	1
2;8.11	10	9	1
2;9.07	44	44	0
2;10.12	21	19	2
3;1.15	11	11	0
3;2.20	15	15	0

Table 4.13 Use of subjects - Paolo

Age	Total inflected verbs	Correct pragmatic use	Incorrect pragmatic use
3;1.27	10	10	0
3;3.23	0	0	0
3;4.25	4	4	0
3;7.10	33	27	6
3;10.19	16	15	1
3;11.17	27	23	4
4;0.29	21	17	4
4;1.28	12	10	2

Table 4.14 Use of subjects - Costanza

Age	Total inflected verbs	Correct pragmatic use	Incorrect pragmatic use
1;1.16	1	1	0
1;12.10	0	0	0
2;2.17	19	19	0
2;4.09	33	33	0
2;6.07	23	23	0
2;7.16	13	13	0
2;9.14	68	68	0

Table 4.15 Use of subjects – Francesca

Age	Total inflected verbs	Correct pragmatic use	Incorrect pragmatic use
2;4.20	3	3	0
2;5.10	0	0	0
2;6.19	9	9	0
2;7.28	10	10	0
2;9.07	8	8	0
2;10.17	15	13	2
3.0.17	26	23	3
3;1.17	58	53	5
3;2.27	66	53	13
3;5.0	17	14	3

Overall, considering all the data, the percentages of correct pragmatic use of subjects are as follows: Costanza 100%, Matelda 96.9%, Francesca 87.3% and Paolo 86.92%.

These results will be used in the final analysis of the weak language in combination with those obtained in the previous sections.

4.10 Combined measurement of the results

The factors analysed so far are rate of acquisition, MLU, vocabulary, production of target-deviant forms, code-switching and discourse pragmatics. Each factor has been analysed on the basis of new or previously employed methods, and has allowed me to point out differences among the four bilingual children and, where the data was available, also between these and other groups of children. The purpose of analysing these factors is to combine the results and determine whether the bilingual children develop Italian as a weak or strong language. Code-switching is excluded since the results do not seem to be significant in this analysis.

The table below (4.16) provides an overview of the data analysed in this chapter. The first observation reported is *stage*. I have divided the data into four stages based on the average MLU in word[42] (Stage 1 = MLU 2-2.2; Stage 2 = MLU 2.3-2.5; Stage 3

[42] A similar type of classification can be found in Serratrice, Sorace and Paoli (2004).

= MLU 2.6-2.8; Stage 4 = MLU 2.9 to 3.1) preceded by the initial of the child's name. The observations were not available for all children (see M4, C1 and P2). Where the MLU value was related to more than one set of data, an average was calculated.

The other observations presented in table 4.16 are the child's age in months, the percentage of articles omitted in obligatory contexts, the amount of vocabulary (calculating the total number of tokens, including adjectives, verbs and nouns), the percentage of target-deviant forms on the total number of utterances and the percentage of pragmatically correct subject realisations. The basic assumption underlying this study is that the combination of the factors analysed would give an indication of weak or strong development. For this reason, the results are assigned numerical values that can be added up.

Table 4.16 Overview of results

Stage	Age	article omission	vocabulary	target-deviant	Subjects
F1	32	57.99%	26.5	3.28%	100%
F2	36	66.66%	44.5	4.22%	88%
F3	39	50.79%	56	5.7%	88%
F4	38	29.03%	108	2.89%	88%
M1	37	0	37	0.72%	90%
M2	38	0	34	0	100%
M3	34	26.05%	36	1.92%	96%
M4	-	-	-	-	-
C1	-	-	-	-	-
C2	30	0	24.5	4.91%	100%
C3	29	0	40	0.96%	100%
C4	28	5.26%	70	0	100%
P1	43	63.33%	31.5	11%	90%
P2	-	-	-	-	-
P3	48	38.23%	50.5	3.59%	86%
P4	46	20%	41	6.33%	84%

These results are quite diverse because they are calculated on the basis of different values and with different methods. Therefore, in order to combine them to provide a final result, I will give each factor a value from 1 to 10. The values are established on the basis of all the data available (from this and also from previous studies). Moreover, these values are not used to create a norm[43] but only to have a uniform measure to compare the children and to give an indication of what could be considered weak and strong development. Ultimately, it would be important to set a norm for each of these factors; however it is necessary to gather more monolingual and bilingual data to standardise this type of analysis. More details on the values assigned to each factor can be found in Appendix G.

[43] I believe that more data is necessary in order to establish a norm.

Table 4.17 Results based on scale

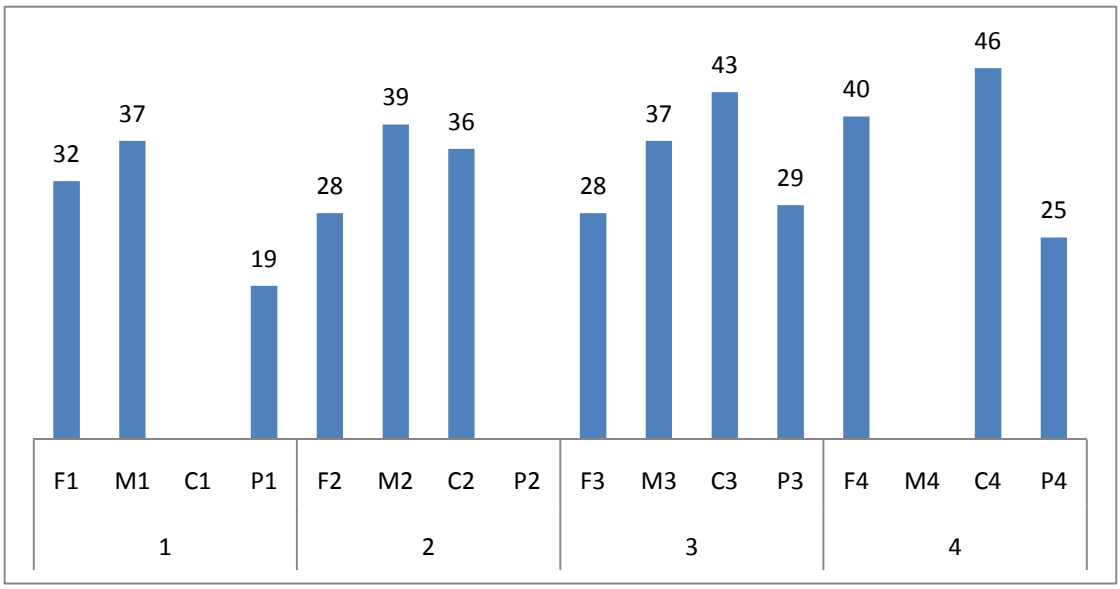
Stage	Age	article omission	vocabulary	target-deviant	Subjects
F1	8	5	2	7	10
F2	6	4	4	6	8
F3	5	5	5	5	8
F4	6	8	10	8	8
M1	6	10	3	10	8
M2	6	10	3	10	10
M3	7	8	3	9	10
M4	-	-	-	-	-
C1	-	-	-	-	-
C2	8	10	2	6	10
C3	9	10	4	10	10
C4	9	10	7	10	10
P1	4	4	3	0	8
P2	-	-	-	-	-
P3	2	7	5	7	8
P4	2	8	4	4	7

If we add up the values obtained at each stage, we obtain the following result:

Table 4.18 Sum of values

F1	F2	F3	F4	M1	M2	M3	M4	C1	C2	C3	C4	P1	P2	P3	P4
32	28	28	40	37	39	37	-	-	36	43	46	19	-	29	25

Figure 4.9 Overview of results



Five factors are examined but four observations are only available for Francesca, while the other children have only three observations. In order to have an equal measurement, I take into account the difference in number of observations and I finally show the results on a scale of 100.

The method can be explained as follows: each observation, for example F1, is connected to a result (F1-32), which is the sum of the results from 5 factors. Each one of these factors can reach a value from 1 to 10, therefore the maximum result for each observation is 50 (5 x 10). Since I have divided the children's data into 4 stages, it follows that the maximum total result for a child who has 4 observations is 200, while for a child who has 3 observations it is 150. Since I aim to have an equal result for all the children based on a final scale of 100, I will add up the results obtained in each observation as in the following example:

F1+F2+F3+F4 = 128

Since this set of data has 4 observations, the maximum total result is 200. In order to have the results on a scale of 100, I divide the result obtained by the sum of observations by 2, as in the following example.

$$\frac{F1+F2+F3+F4}{50+50+50+50} = \frac{128}{200} : 2 = \frac{64}{100}$$

The other three children only present three observations, therefore the measurement follows the same criteria but considering that the maximum total can be 150.

$$\frac{M1+M2+M3}{50+50+50} = \frac{113}{150} : 1.5 = \frac{75.3}{100}$$

4.11 The weak language scale

The purpose of adding up the factors was to obtain a homogeneous result that could allow us to compare the children. The final results are used to determine whether Italian is a weak or strong language. The main limitation of this analysis is that the values and therefore the scale are based on a limited amount of data, and not on normative standards. However, this could be the starting point for further research and for the collection of larger data samples[44].

The method I employed for the assessment of the weak language takes into account values that have not been considered in combination in previous studies on the weaker language. It accounts for 5 factors that are analysed on the basis of different phases of development. The preliminary assumption is that a combination of a slow acquisition rate, a limited vocabulary, a high number of word order and morphology errors and also a pragmatically incorrect selection of null vs. overt subjects can be regarded as an indication of weak language development[45]. Another limitation of the assignment of values is the significance of each factor. We may find that one or more of the factors analysed might be more significant than others in the analysis of the weak language. Again, the limited amount of data used in this study does not allow a reliable statistical testing of the significance of each factor.

The main purpose of this study is to set criteria to determine whether it is possible – and to what extent – a bilingual child is developing Italian as a weak or strong language. Even given the limitations described above, this analysis constitutes an important contribution to the study of the weak and strong development of Italian, and it can be seen as a starting point for future research in this area.

[44] Recently, a new collaborative project has been launched to fill this gap in the data on the acquisition of Italian in different contexts. Details of this project, called *La ricerca fondamentale sul linguaggio al servizio della lingua italiana: documentazione, acquisizione monolingue, bilingue e L2, e ideazione di prodotti multimediali. - progetto FIRB (2008)* can be found on http://www.ciscl.unisi.it/ricerca.htm.

[45] It has to be taken into account that there are different degrees of weakness. By using the Weak Language Scale, it should be possible to determine the "degree of weakness" of a language.

In order to determine whether a language is strong or weak, I establish the following criteria[46]:

- A result between 70 and 100 corresponds to strong development, and it is subdivided into three levels (*S[47] I, S II and S III*)
- A result over 40 and below 70 corresponds to weak development, and it is subdivided into three levels (*W I, W II and W III*)
- A result below 40 corresponds to a considerably weak development

Figure 4.10 The Weak Language Scale

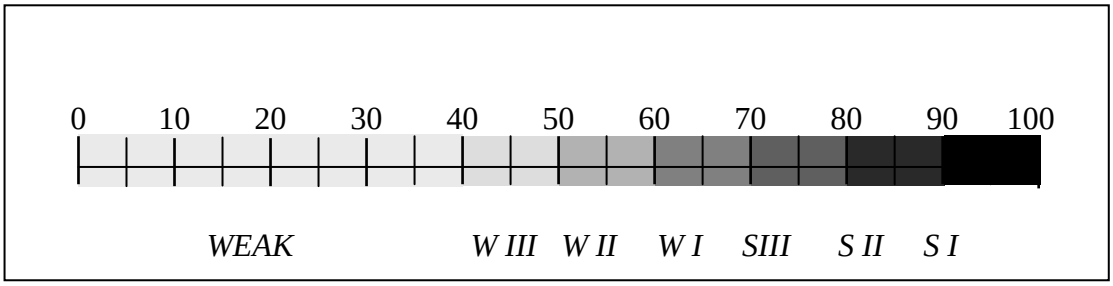
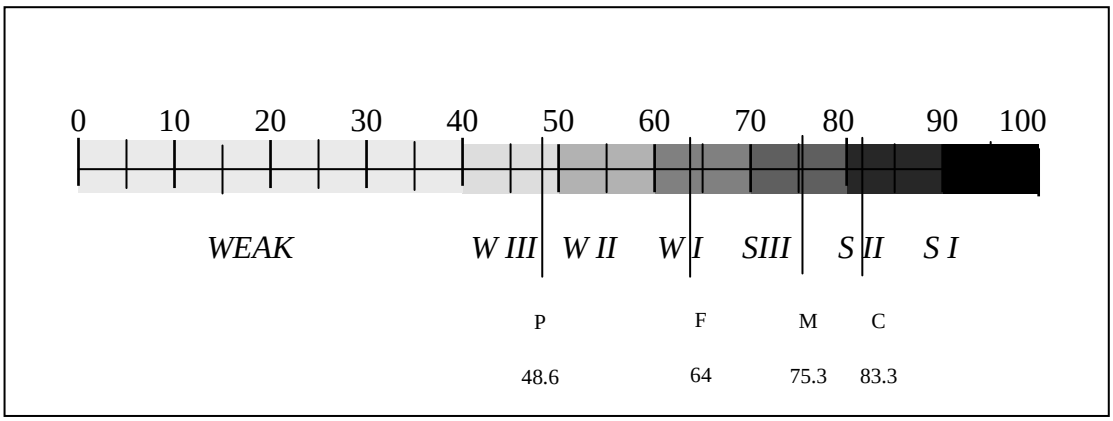


Figure 4.11 Results on the basis of the Weak Language Scale



According to this scale, Costanza and Matelda develop Italian as a strong language, while Paolo and Francesca as a weak language. These results are consistent with the hypotheses formulated in chapter 3. As the various analyses presented throughout this chapter have shown, Costanza develops more similarly to a monolingual Italian child than the other children, and this result emerges also in the final assessment.

[46] I calculated that a monolingual child would score a result over 70, therefore I assumed that 70 could be used as a threshold of strong language development. This criterion is quite arbitrary and more data is needed to support the validity of this measurement.

[47] S stands for Strong, W stands for Weak.

Both Paolo and Francesca develop Italian as a weak language, but Francesca reaches a higher score, which shows that her Italian is less weak than Paolo's. The advantage of having a Weak Language Scale lies in the fact that it is possible not only to compare data from different children, but also to determine to what extent the language is weak or strong. The subdivision of weak and strong into 3 levels (S I, S II, S III and W I, W II, W III) is used to further highlight the fact that weakness is not a static phenomenon and also to account for the different degrees of weak or strong linguistic development.

The Weak Language Scale has been used to establish criteria and numerical values to assess a bilingual child's production and it has been built on the basis of Italian data. However, the same method can be applied to other languages. Due to the lack of data, it is not possible to further confirm the validity this scale. However, a further validation of the reliability of this method can be verified by applying it to data from other bilingual children.

4.12 Chapter summary and conclusion

In this chapter I have introduced the methodology used to analyse the development of Italian in order to determine whether it is a weak or strong language. I have also clarified the crucial difference between this and previous studies on weak language development. The main difference is the type of data analysed: while previous studies aim to determine whether one language is weaker than the other within an individual, in this thesis I only take into account Italian data. Another important aim of this chapter is to select criteria to determine whether Italian is a strong or weak language. The factors analysed are rate of acquisition, vocabulary, number of target deviant forms, MLU and discourse pragmatics. I analysed these factors separately and then in combination in order to obtain a single result that can be used to compare the four children. The final results show that Costanza and Matelda develop Italian as a strong language, while Paolo and Francesca as a weak one. These results, which are presented through the Weak Language Scale, also allow us to observe to what extent the children develop Italian as a strong or weak language.

One of the main questions initially addressed in this thesis concerns the nature of the weak language. As it has been explained in chapter 2, there is not yet agreement on

what a weak language is and how weakness should be assessed. My first aim was to determine whether the factors used to analyse the weak language were significant. On the basis of the analysis of bilingual and monolingual Italian data, I found that all the factors I examined showed significant differences in the results obtained by the different children. However, I decided to exclude the analysis of code-switching because it didn't show significant results.

It emerges from my analysis that Italian develops as a weak language if the child's language exhibits the following characteristics:

- Slow rate of acquisition
- Short MLU
- Presence of a high number of word order errors
- Presence of a high number of agreement errors
- Presence of a high number of verb inflection errors
- Limited lexicon
- Difficulty making pragmatically correct choices in the use of overt and null subjects.

These characteristics can affect the linguistic development at different stages and to a varying degree, but if they are consistently found they can suggest that the child's Italian is weak.

An important result emerging from this analysis is that bilingual children perform differently from each other and some perform in a more native-like fashion than others. There are however limitations to this type of analysis related to the lack of large samples of monolingual and bilingual data, which would allow a more reliable testing of each factor. More data needs to be used to analyse each factor, to compare the children, and to validate the numerical values used in the weak language scale. The availability of larger sets of data would also allow us to standardise such a test. However, the methodology employed can constitute an initial step towards a comprehensive assessment of the weak language, which can be used for Italian and also adapted to other languages.

In the next chapter, I will look at the children's use of subject inversion, a structure at the interface between syntax and pragmatics, with the aim to determine whether its analysis can be complementary to the study of the weak language.

CHAPTER 5

The weak language and the syntax-pragmatics interface

5.1 Introduction

In chapter 4, I presented an analysis of the weak language based on five factors, namely MLU, lexicon, rate of acquisition, production of target-deviant forms and discourse pragmatics. Since research has shown that interface properties can represent a difficulty for bilinguals (Müller 2008, Belletti and Leonini 2004), I aim to further test weakness by examining in more detail the use of subjects and specifically the production of subject inversion. As I explained in chapters 1 and 2, in Italian it is possible to find SV or VS word order, on the basis of syntactic, pragmatic and lexical constraints. When the subject is preverbal it represents the topic (old information), when it is in postverbal position it represents focus (new information). The distinction between topic and focus depends on the shared knowledge among the speakers, as shown in the following examples[48]:

5.1 Che è successo?
What happened?
È crollato un palazzo. (VS)
Un palazzo è crollato. (SV)
A building collapsed

5.2 Sai dove sono le chiavi?
Do you know where the keys are?
Le ha prese Gianni (VS)
Gianni le ha prese (SV)
Gianni took them

In both examples there is a question followed by an answer in which the subject constitutes new information. Therefore, although both VS and SV are grammatically possible, the two sentences require the placement of the subject in a

[48] Example 5.1 is from Belletti, Bennati and Sorace (2007, p. 665)

postverbal position in order to put an emphasis on the fact that the information is new to the hearer.

As I stated in chapter 2, there are two main reasons for choosing to analyse subject inversion. Firstly, the word order Verb-Subject (VS) is frequently used in Italian but not in English[49] (this does not constitute an argument for cross-linguistic influence but rather for the fact that children have to rely on their syntactic and pragmatic ability in Italian). Moreover, research has shown that the VS order is a "fragile" structure because it involves the simultaneous activation of syntactic and pragmatic knowledge and therefore requires a heavy processing load (Belletti and Leonini 2004: 26). In addition, there is evidence demonstrating that non-native and near-native speakers of Italian can use null and overt subjects appropriately, but they have difficulty in mastering subject inversion (Belletti, Bennati and Sorace 2007, Belletti and Leonini 2004, Bettoni, Di Biase and Nuzzo 2009). On the basis of these findings, it seems appropriate to test subject inversion as a further factor to assess weakness. The prediction is that children who develop Italian as a weak language have difficulty mastering structures involving the simultaneous activation of two different domains. This prediction will be tested by examining spontaneous and experimental data on the production of subject inversion. If the results show non-native-like performance, two explanations can be considered: firstly, it could be argued that in cases of exposure to limited input, children might not come across many utterances containing VS order (VS is generally less frequent than SV order). It is also possible that the input is impoverished because of cross-linguistic influence or attrition in the adult's production. To rule out this hypothesis and to observe the presence of postverbal subjects in the input, I will also examine the adult data. The second possible cause for a non-native-like performance could be the nature of the subject inversion: the complexity of the interface between syntax and pragmatics can lead some children to make word order errors or to avoid the structure. The testing of these hypotheses will provide a contribution to the understanding of the relationship between weakness, interface processing and the role of the input.

In the analysis presented in this chapter, I use both longitudinal and experimental data. The main reason for combining two analyses is the availability of samples.

[49] Stylistic inversion can be found in English, but it is not typical of child language.

The longitudinal data provides very few examples of postverbal subjects, therefore I will also test the children's production of subject inversion though two elicitation tasks. Both tasks aim to trigger the production of postverbal subjects, however the second one is more complex, since it also requires the production of direct object pronouns[50].

Finally, it has to be considered that in terms of word order, English provides consistent evidence of the preverbal position of the subject and it presents a structure which is less complex than Italian. This could give ground to the hypothesis that children may make word order errors because they rely on the knowledge of a simpler and more frequent word order (SV) which is available both in Italian and English (Müller and Hulk 2001, Serratrice, Sorace and Paoli 2004). However, my hypothesis is that weakness rather than cross-linguistic influence should be regarded as the cause of the difficulty of processing interface structures[51].

5.2 Methodology

In this chapter I present two sets of data. The first consists in the spontaneous longitudinal data from the four bilingual Italian-English children (see chapter 3). The second consists in experimental data collected from the same children and from other control groups. The purpose of analysing the spontaneous data is to

[50] The presence of direct object pronouns in the experiment design is not accidental. Direct (5.3) and indirect (5.4) object pronouns are another example of the interface between syntax and pragmatics and this will be expected to present a similar challenge to subject inversion. From a syntactic point of view, object pronouns in Italian occupy a preverbal position, and they agree in number and gender with the object they refer to. From a pragmatic point of view, they are used to refer to an antecedent, and they 'appear in order to mark the dislocated noun phrase as presupposed either in the discourse or by the hearer' (Müller 2008: 73). As Müller argues, when the presupposed or known object is present, the use of an object pronoun is obligatory, as shown in the two examples below.

5.3 Compro il giornale e **lo** leggo in autobus.
buy1sg the paper(m) and **it**(m) read1sg on bus *I buy the paper and I read it on the bus.*

5.4 Chiamo Maria e **le** dico che sono in ritardo.
ring1sg Mary(f) and **to her**(f) say1sg that be1sg late. *I ring Mary and I tell her that I am late.*

According to Müller and Hulk's (2001) prediction, bilingual children who acquire a language that requires object pronouns and one that allows null objects will omit the object in contexts where it is required. Müller (2008) suggests that bilinguals might not produce object pronouns before age 4. Even though English does not allow object drop like German, it is possible to assume, following the claim made by Hulk and Müller (2000) that if English presents a structure that is less complex than Italian, the children will use the less complex structure in both languages. This observation is related to the ability of the children to compute complex structures which require a high processing load (Müller 2008).

[51] It could be assumed that weakness might also cause cross-linguistic influence. However, the discussion of this hypothesis goes beyond the scope of my research.

observe the distribution of subjects in Italian. The analysis of subject distribution replicates the method employed by Lorusso, Caprin and Guasti (2004) and Cabré Sans and Gavarró (2007), and it provides additional results on the use of postverbal subjects. The experimental data is used to further assess the children's ability to produce postverbal subjects in a controlled task. The bilingual children were tested on the basis of a repetition task and an elicitation task. More details on the tasks and the results are provided in sections 5.5 to *5.5.3*. The results of the second task are compared to those of monolingual and bilingual control groups. Finally, the results are considered in the light of the analysis of the weak language.

5.3 Subject distribution

Many recent studies have analysed the acquisition of subjects in Italian and other null subject languages from a generative perspective. One of the first studies was carried out by Hyams (1986), who showed that all children initially produce a high rate of null subjects, and later set (or keep) the language-specific parameter value. Since then, many studies have focused on subject omission across languages[52], and in recent years there has been an increasing interest in the discourse and pragmatic constraints determining the distribution of null and overt subjects both in child L1 and adult L2 acquisition. In an analysis of subject distribution in child Italian, Serratrice (2005) found that from the earliest stages (as early as MLU 2.0) children make the correct pragmatic choice and rarely produce referentially ambiguous sentences.

Studies on bilingual language acquisition can offer a further insight into the issue of subject distribution. The children involved in the research for this thesis are English-Italian bilinguals. Since English is not a null-subject language, it is possible to assume that cross-linguistic influence might be found in the spontaneous data. Therefore, it is possible to hypothesise that children might have difficulty in the selection of null or overt subjects and produce ungrammatical sentences in terms of informativeness. Paradis and Navarro (2003) provide evidence for this claim, by showing that English-Spanish bilingual children produce more overt subjects than monolinguals, often in cases

[52] For a review of studies on the acquisition of null subjects, see Guasti (2007).

where they are not required. Serratrice et al. (2004) compared the data from Carlo to other monolingual Italian data and found that at all stages he produces more overt subjects than monolingual children. This finding can be considered evidence of cross-linguistic influence (see also Serratrice 2002). However, cross-linguistic influence is not always found in bilinguals acquiring a null and a non-null subject language: in a study of 2 German-Spanish bilingual children (aged respectively 1;7-3;3 and 2;01-3;05), Hinzelin (2003) found no instances of cross-linguistic influence. A more frequent production of overt subjects has also been attributed to the input. Paradis and Navarro (2002) analysed the parents' data, calculating the percentage of overt and null subjects, and found a correlation between the overuse of overt subjects in the children's data and in the adults' input. As the authors observe, it is not possible to understand from the data whether the non-target use of subjects, which seems to be an effect of cross-linguistic influence, depends on performance factors related to the child, on the input variety or on the interplay of the two.

In what follows, I present the analysis of the use of subjects among the bilingual children. In this analysis, mostly based on Lorusso, Caprin and Guasti (2004), all subject forms are considered, including pronouns, demonstratives, proper names, full and bare NPs and quantifiers (see also Serratrice 2005). The type of verb is also considered, following the distinction among transitive, unaccusative and unergative verbs. Only declarative sentences with inflected verbs are included in the analysis. These criteria match those established by Lorusso, Caprin and Guasti (2004)[53], making the two studies comparable.

Table 5.1 shows the distribution of null and overt subjects. Lorusso et al.'s study shows that monolingual children produce mostly null subjects (67% to 79%). Matelda and Costanza produce respectively 75% and 66% null subjects, while Paolo and Francesca show a different pattern from monolinguals. In particular, Francesca uses a high number of overt subjects in comparison to other bilinguals and monolinguals. The data is presented in percentages as in Lorusso et al.'s study.

[53] From this point, this study will be referred to as Lorusso et al's.

Table 5.1 Null vs. Overt Subjects

	Null Subjects	Overt Subjects
Francesca	32%	68%
Matelda	75%	25%
Paolo	51%	49%
Costanza	66%	34%
Children's Average	56%	44%

Another important factor is the analysis of the type of verbs in the sentence (see Table 5.2), because the distribution of preverbal and postverbal subjects is related to the argument structure of the verb (Pinto 1997). Monolingual children produce mostly overt subjects with unaccusatives (32% to 41%), and almost in equal measure with unergatives (25%) and transitives (22%). Child and adult bilinguals show similar results.

Table 5.2 Overt subject distribution across verb classes

	Unaccusatives	Unergatives	Transitives
Francesca	52%	4.6 %	43.4%
Matelda	34.7%	4.3%	60.8%
Paolo	51.51%	1.49%	47%
Costanza	41%	14%	45%

An interesting result is provided in table 5.3, which shows the distribution of overt subjects. The monolingual children analysed by Lorusso et al. produce about 62% preverbal and 38% postverbal subjects. Again, this result closely matches Costanza's and Matelda's. Paolo and Francesca exhibit a different trend, producing a very limited number of postverbal subjects. A further analysis is provided in table 5.4, which shows the distribution of preverbal and postverbal subjects across verb classes.

Table 5.3 Overt preverbal vs. overt postverbal

	Overt Preverbal	Overt Postverbal
Francesca	93.8%	6.2%
Matelda	61.5%	38.4%
Paolo	92.4%	7.5%
Costanza	53%	47%

Table 5.4 Overt subject position across verb classes: percentages

	Unaccusatives		Unergatives		Transitives	
Francesca	Preverbal 98.5%	Postverbal 0.5%	Preverbal 83.3%	Postverbal 16.6%	Preverbal 89.2%	Postverbal 10.8%
Matelda	Preverbal 50%	Postverbal 50%	Preverbal 0%	Postverbal 100%	Preverbal 64.28%	Postverbal 35.72%
Paolo	Preverbal 91%	Postverbal 9%	Preverbal 100%	Postverbal 0%	Preverbal 93.54%	Postverbal 6.46%
Costanza	Preverbal 66.6%	Postverbal 33.4%	Preverbal 57%	Postverbal 43%	Preverbal 39%	Postverbal 61%

Lorusso et al.'s data show that the highest percentage of verbs occurring in postverbal position and unaccusatives, both in the child and the adult data. The bilingual data shows a different pattern. Only Matelda shows an equal use of unaccusatives in preverbal and postverbal position. However, as the number of examples in her data is very limited, I would not consider this evidence sufficient to make any generalisation. The highest percentage of postverbal subjects occurs with transitive and unaccusative verbs, unlike in the monolingual data. It emerges from this analysis that monolingual children exhibit a uniform distributional pattern, while bilinguals show more variation. However, overall Matelda and Costanza seem to show more similar subject distribution to monolinguals. Another interesting result is the limited use of postverbal subjects with unaccusative verbs. For this reason, in the section that follows I will present the data by using the number of tokens rather than the percentages.

5.4 Postverbal subjects

In this section, I provide an overview of the distribution of postverbal subjects in the spontaneous bilingual data. A low number of postverbal subjects emerge from the longitudinal data. Matelda and Francesca produce 8 postverbal subjects, Paolo 6 and Costanza 24. Table 5.5 shows the distribution of postverbal subjects on the basis of the verb type. Francesca, Matelda and Costanza use postverbal subjects mostly with transitive verbs, while Paolo mostly uses them with unaccusatives.

Table 5.5 Postverbal subjects classified by verb types

	unaccusative	unergative	transitive
Francesca	1	1	6
Matelda	2	1	5
Paolo	3	1	2
Costanza	7	3	14

As I show in Table 5.6, the majority of postverbal subjects are nouns and Costanza is the only child who uses all types of subjects. Sentences 5.3 and 5.4 are examples of postverbal structures from the spontaneous data.

5.3 *MAT: Viene il mio amichetto.
My friend comes.

5.4 *COS: Piange il papà.
The dad cries.

Table 5.6 Types of postverbal subjects

	pronoun	noun	proper noun	demonstrative
Francesca	0	6	1	1
Matelda	2	5	1	0
Paolo	0	6	0	0
Costanza	1	11	7	5

In terms of function of inversion, most of the postverbal subjects that emerge represent presentational focus (e.g. *é caduta la banana – the banana has fallen*).

There are no instances of locative inversion. Costanza and Matelda are the only children who produce structures with presentational, contrastive and narrative function[54].

The evidence presented in this section shows that postverbal subjects are used mostly with transitive verbs and the subject is generally a noun phrase. The bilingual children differ from the monolingual ones mainly in the use of verbs. It also emerges that most inversion structures have a presentational function. The main finding emerging from the analysis of the bilingual longitudinal data is that the occurrence of postverbal subjects is quite limited. Costanza is the only child who produces postverbal subjects almost in every recording, with different types of verbs and subjects, resembling the behaviour of monolingual children and adults.

Since the number of postverbal subjects is quite limited, the question rises whether this can be attributed to performance limitations, cross-linguistic influence or just to the fact that during the recording sessions the discourse-new information was simply not required. Cross-linguistic influence could justify the absence of subject inversion structures, if we assume that this influence manifests itself in the avoidance of the structure. However, there is a lack of clear examples of cross-linguistic influence in the data. The only example that I could relate to this phenomenon is the following sentence, where we could expect to find the subject "mamma" in the postverbal position:

5.5 *INV: Chi te l'ha comprato questo?
Who bought you this?
*FRA: Mamma l'ha comprato.
Mum bought it.

The adult data also shows that the parents who provide a higher qualitative input (see chapter 6) produce more postverbal subjects. However, even in the adult data the samples are scarce. This could be interpreted as a cause for the lack of postverbal subjects in the child data. However, the data from Costanza challenges

[54] Presentational focus has the role to put an emphasis on the constituent that represents new information (e.g. *è arrivato **il treno** – the train has arrived*). Contrastive focus has the same function, but it also stresses the opposition of two elements (e.g. *ha vinto l'Italia, non la Francia – It was Italy who won, not France*). Narrative inversion is used generally when quoting somebody else's words (e.g. *"Ti mangerò", disse il lupo – "I will eat you", said the wolf*).

this hypothesis, since the girl produces a high number of postverbal subjects, and her mother does not. Therefore, I would like to argue that the scarcity of postverbal subjects in the children's data should not be attributed to the lack of production by the adults, but rather to the overall lack of input. This hypothesis is supported by the results presented in chapter 6, which show a correlation between the input and the quantity of postverbal subjects produced by the children.

Since the evidence found in the spontaneous data is too scarce to carry out a reliable analysis of the children's production of postverbal subjects, I test their performance in two elicitation tasks.

5.5 The elicitation tasks

As the spontaneous data presented so far shows, postverbal subjects are less frequent than null or overt preverbal subjects and they are overall very infrequent. In order to achieve a better understanding of the production of postverbal subjects, two elicitation tasks were designed. The first aims at testing the production of postverbal subjects with unaccusative verbs and the second with transitive verbs. The latter also requires the use of direct object pronouns and it is therefore a task that involves a higher processing load, as I will show in the next section.

5.5.1 Pimpa

In this task the investigator reads a short story to the child. The story is made of 10 short sentences corresponding to a picture, each presented on a separate page. The format is similar to a book with 10 pages. The main character of the story is Pimpa, a character all children are familiar with[55]. Two of the pages in the book contain movable objects and characters that the children can play with. One is a penguin that can be moved on the page, the other is a pair of clouds that can be moved in the sky and show the sun behind them. In the first part of the task the investigator reads the story to the child, and shows him/her how to move the objects. First, the objects are moved and then the sentence is pronounced. The children enjoyed doing this type of activity and "interacting" with the book. The task should elicit two sentences which contain a presentational inversion structure

[55] I chose Pimpa because I knew that all the children had already read stories involving Pimpa and they were familiar with the characters in those stories.

with unaccusative verbs. The first is '*Guarda! È uscito il sole!*' (*Look, the sun has come out!*), and the other is '*Guarda! È arrivata Nina la pinguina!*' (*Look, Nina the penguin has come!*). Figures 5.1-5.3 show the picture the children are presented with and the way it changes after the children move the objects.

Figure 5.1 Elicitation task phase 1: the clouds cover the sun.

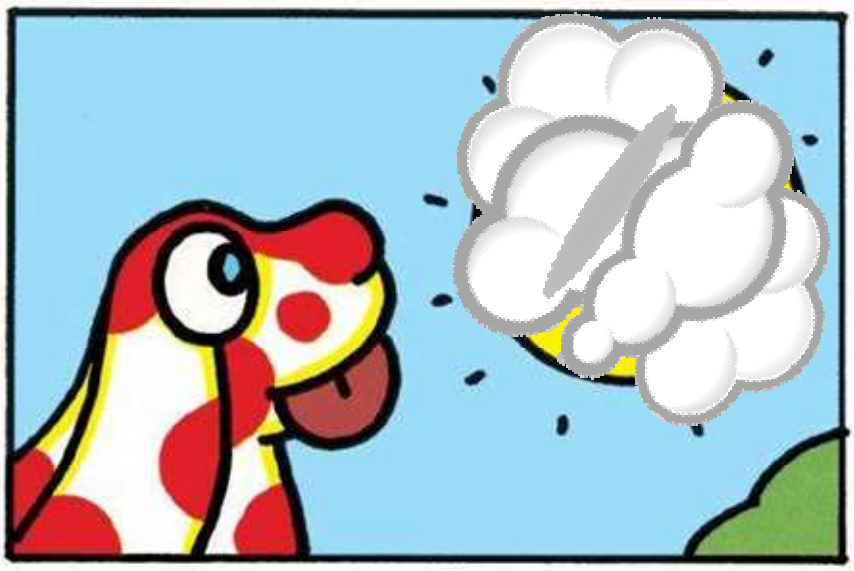


Figure 5.2 Elicitation task phase 2: the child moves the clouds.

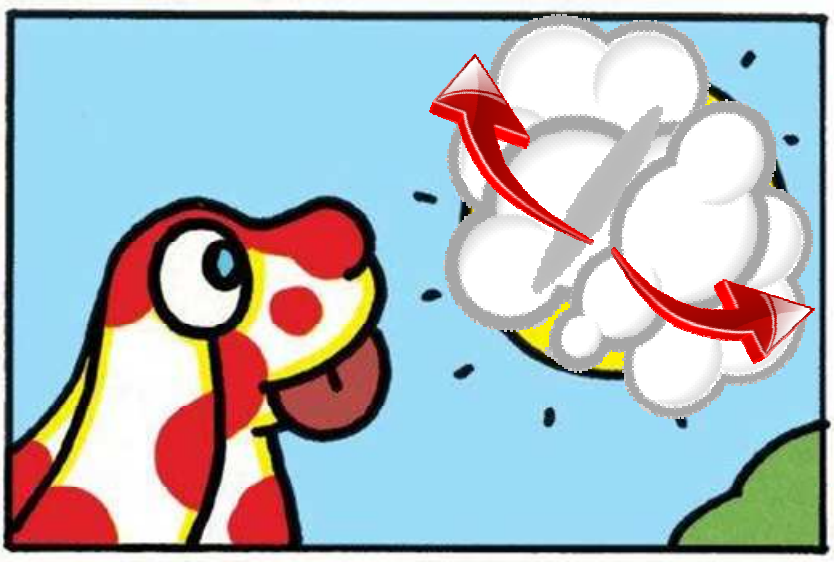
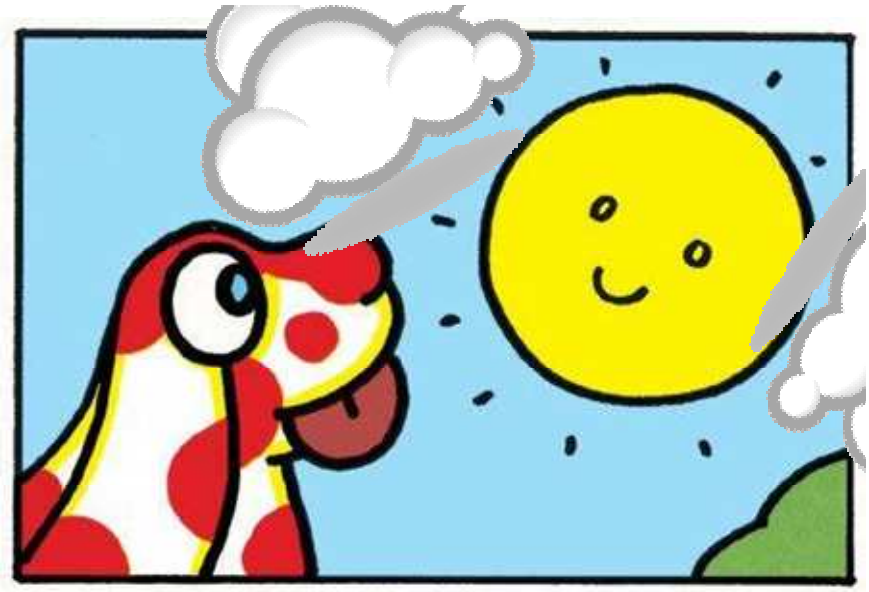


Figure 5.3 Elicitation task phase 3: the child sees the sun.



The target answer is '*È uscito il sole!*'. However, other possible answers can be considered valid. What has to be constant is the emergence of new information, which is conveyed by a postverbal subject. None of the children produced an ungrammatical sentence (table 5.7). However, Matelda and Costanza use verbs that are more appropriate to communicate the change of state.

Table 5.7 Results

Child	Response	Verb Type	Subj. position
Costanza	è uscito il sole! *the sun has come out!*	unaccusative	postverbal
Matelda	arriva il sole! *the sun comes!*	unaccusative	postverbal
Francesca	c'è il sole! *there is the sun!*	copula/unacc.	postverbal
Paolo	il sole! *the sun!*	-	-

In the second elicited sentence, none of the children produced the target answer, and the only responses they produced included the name "Nina" or "Nina la pinguina". I think this result is related to the nature of the task: while in the first example (the one with the sun and the clouds) the children are involved in the action that produces the change of state, in the second one the children already see the penguin on the page and even though they have to move it onto the scene, they find it harder to relate to the whole scenario.

This task was used to test the ability of the children in producing a sentence with new information and also to test the potential of this type of task. Even though the children seemed to enjoy the story, the task was too long and it was difficult to keep their attention focused on more than three or four pictures. It emerges from the results that the two children who, according to the weak language scale, are developing Italian more closely to an L1 are able to use the appropriate information structure with unaccusative verbs. These results will be discussed in more depth in section 5.6.

5.5.2 Animali in cucina[56]

The second elicitation task is more complex and it tests the production of postverbal subjects in more complex sentences. The task was perfomed by the four children twice, first during the last recording session and also after about 8-9 months. I proposed the test to the same children after a long period in order to determine whether their performance improved with time[57].

The task requires the children to answer four questions while they watch a short cartoon. The aim is to trigger responses with postverbal subjects by asking questions starting with '*che fine ha fatto* ...', which can be translated as '*what happened to*...'. This type of question requires a response in the form of a full answer. A sentence such as "who had the apple" could just be answered using the subject "the spider", while a question such as "what happened to the apple/where's the apple gone?" requires the use of a longer response which also contains a verb. While the first elicitation task involved the use of unaccusative verbs, the second one requires the use of transitive verbs.

[56] Animali in cucina (animals in the kitchen) is the title given to the cartoon.
[57] The children did not seem to remember the cartoon the second time. They also showed no signs of training effect.

5.5.2.1 Participants and procedure

This experiment aims at testing the children who were involved in the longitudinal study (see chapter 3). In addition, the following four control groups were tested:

- GROUP A: 6 monolingual Italian children - aged 3, 4 and 5
  (two participants for each age group)
- GROUP B: 7 monolingual Italian children - age range 2;0-2;10
- GROUP C: 10 monolingual Italian adults - age range 19-60
- GROUP D: 10 Italian-English bilingual adults[58] - age range 24-35

The participants are presented with a cartoon shown on a computer screen through PowerPoint slideshow. The transition from a slide to the next is not automatic, in order to give the children enough time to answer (time is not considered a factor in the children's response). The next slide is presented after the child answers the question. The investigator reads the story, asks the questions and controls the slide transition. The experiment is audio-recorded and the answers are transcribed. The cartoon is made of 13 slides, each one containing an animation, in order to make the story more attractive and to focus the child's attention. Each slide corresponds to a sentence, which is read to the child. There are four questions that are directly addressed to the child. The experiment lasted a maximum of 50-60 seconds with children and 30-40 with adults.

The cartoon is designed as follows: the inital scene shows a table with a drink, an apple, a bowl of cereal and some carrots. Each time, an animal comes in from an open window and eats or drinks one of the items on the table. After the animal has disappeared, the mother appears in the scene and poses a question addressed to the child (using his/her name, and asking what happened to the item ('Francesca, che fine ha fatto la mela?', '*Francesca, what happened to the apple?*'). The child is then shown an animation related to the answer. It has to be noted that the animation shows the animal appearing in the room and eating or drinking. The figures below show some stills of the animation related to one of the questions (the full details on the experiments can be found in Appendix H).

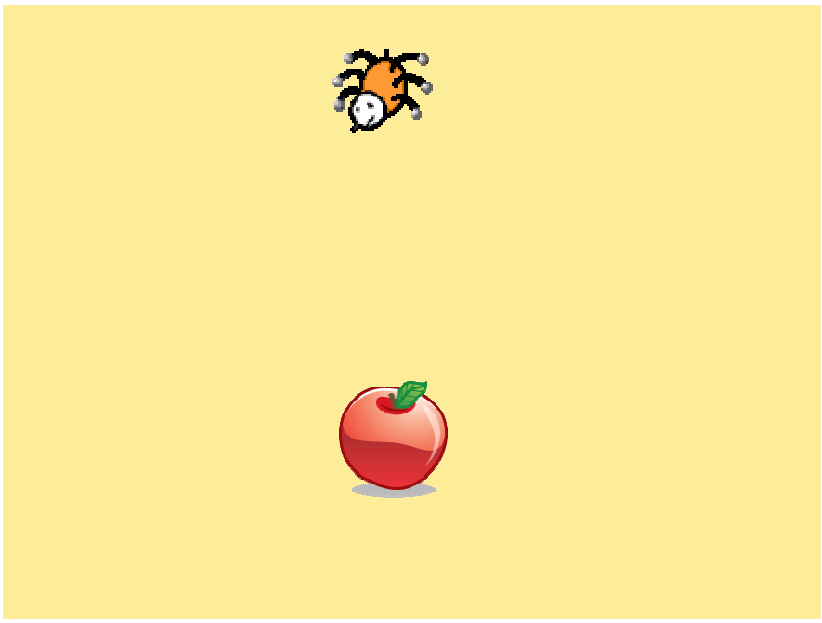
[58] The adults tested have been living in Ireland for at least 3 years.

Figure 5.4 Initial scene

Figure 5.5 The spider comes in and eats the apple.

Figure 5.6 The mother comes and says ‘Che fine ha fatto la mela?’

Figure 5.7 The fourth slide shows the spider eating the apple.



The target answers should contain a postverbal subject like in the following sentence:

5.6 A: Che fine ha fatto la mela?
What happened to the apple?
B: L'ha mangiata il ragno.
It has eaten the spider
The spider ate it.

As we see in this example, the structure of the target sentence includes a direct object pronoun that refers to the referent expressed in the question, followed by the past tense of the verb *to eat*. Both the pronoun and the participle form of the verb agree with the gender and number of the object. The verb is followed by the subject, which is in a postverbal position, representing new information.

In designing the task, the complexity of the sentence was taken into account. Each question requires an answer similar to 5.6 B, but the objects it refers to are of different gender and number. Even though the questions require the same type of answer (Pron+Verb+Subject), they are quite complex, because the pronoun and the verb have to agree in gender and number with the antecedent referent. Therefore, from a processing point of view, these answers require a high processing load, since they require the use of two structures at the interface between syntax and discourse-pragmatics (reference to an antecedent and subject inversion).

Table 5.8 shows for each object the gender, number and the corresponding pronoun and verb agreement.

Table 5.8 Morphological construction of the sentences in the task

Noun	Gender	Number	Direct object pronoun	Verb
1. mela (apple)	F	S	l'/la	mangiat**a**
2. cereali (ceral)	M	P	li	mangiat**i**
3. carote (carrots)	F	P	le	mangiat**e**
4. succo di frutta (juice)	M	S	l'/lo	bevut**o**

It is also possible to use the verbs *mangiare* and *bere* in their reflexive form *mangiarsi* and *bersi* (e.g. se l'è mangiata il ragno/l'ha mangiata il ragno). This would not change the dynamic of the sentence. The main difference between mangiare and mangiarsi is the auxiliary selection: while mangiare requires the auxiliary *avere* (*to have*), the reflexive mangiarsi requires *essere* (*to be*).

5.5.3 Results

The principal aim of the test is to determine whether the children produce postverbal subjects to convey new information[59]. The 10 bilingual and the 10 monolingual adult controls produced the target structure (Pronoun-Verb-Subject). The older monolingual children (GROUP A) also produced the same answers as the adults. The younger monolingual children (GROUP B) produced shorter answers, as shown below (table 5.9-5.10). Table 5.12 shows the verb agreement (e.g. mangiata/ha mangiata referring to mela). Correct agreement is marked with an x.

Table 5.9 Word order – monolingual children GROUP B

Question	Child 1 age 2;10	Child 2 age 2;0	Child 3 age 2;4	Child 4 age 2;1	Child 5 age 2;0	Child 6 age 2;4	Child 7 age 2;0
1.	Pron-V-S	S	Pron -V	S	Pron -V-S	Pron -V-S	S
2.	Pron -V-S	-	-	S	Pron -V-S	S	S
3.	Pron –V	O	-	-	-	S-V-O	S
4.	Pron -V-O	V-O	-	-	Pron -V	S	Pron -V

[59] From a developmental perspective, there are universal properties that govern the acquisition of direct object pronouns and other clitics. In all languages object pronouns do not appear in the wrong position, and they tend to be omitted in obligatory contexts until about age 3 (Tsakali and Wexler 2003). Morevoer, according to Tsakali and Wexler (2003) children are aware that clitics have to agree with the participle. Cipriani et al. (1993) show that direct object pronouns are the first pronouns to appear between 20 and 26 months, but they are produced in more than 75% of obligatory contexts between 26 and 34 months.

Table 5.10 Agreement – monolingual children GROUP B

Question	child 1 age 2;10	child 2 age 2;0	child 3 age 2;4	child 4 age 2;1	child 5 age 2;0	child 6 age 2;4	child 7 age 2;0
1.	x	-	x	-	x	x	-
2.	x	-	-	-	x	-	-
3.	x	-	-	-	-	-	-
4.	x	-	-	-	x	-	x

Tables 5.11 and 5.12 show the bilingual children's results.

Table 5.11 Word order – bilingual children

Question	Matelda age 3;2	Francesca age 3;6	Costanza age 2;9	Paolo age 4;1
1.	V	-	Pron-V-S	-
2.	Pron-V	-	Pron-V-S	S
3.	V	V	Pron-V-S	V
4.	Pron-V	S	S-V-O	S

Table 5.12 Verb Agreement – bilingual children

Question	Matelda age 3;2	Francesca age 3;6	Costanza age 2;9	Paolo age 4;1
1.	x	-	x	-
2.	-	-	-	-
3.	x	-	x	-
4.	x	-	x	-

The results of the task show that monolinguals are able to produce the target answer as early as age 2;0. However, of the 7 children tested, 4 did not produce the target answer in any of their responses. Moreover, all of the target answers not only show correct word order, but also correct agreement.

Of the bilingual children, only Costanza produces the target word order, also with the correct agreement in three responses. Francesca and Paolo only produce either verbs or subjects, while Matelda produces two object pronouns and also three correct verbs, but no subjects.

The test was performed a second time after 8-9 months. The reason for replicating the experiment was to see whether the advancement in linguistic development would produce more target-like results.

Table 5.13 Results 2: Word order

Question	Matelda age 3;10	Francesca age 4;1	Costanza age 3;7	Paolo age 4;10
1.	V-S	S-V	Cl-V-S	S-Cl-V
2.	V-S	S-V	Cl-V-S	S-Cl-V
3.	V-S	S-V	Cl-V-S	S- V
4.	S	S-V	Cl-V-S	S- V

Table 5.14 Results 2: Verb Agreement

Question	Matelda age 3;10	Francesca age 4;1	Costanza age 3;7	Paolo age 4;10
1.	-	-	x	-
2.	-	-	x	-
3.	-	-	x	-
4.	x	-	x	-

The results from the experiment run the second time are clearer firstly because all children give a response. Costanza's performance improves and she produces four target sentences. Matelda produces postverbal subjects, but she omits pronouns and maybe this causes the lack of verb agreement. Paolo and Francesca only produce preverbal subjects. These results show that Matelda and Costanza, the children who develop Italian as the strong language, are aware of the pragmatic constraints in the production of new information structure. The fact that Paolo and Francesca have not yet mastered this structure respectively at age 4;10 and 4;1 is significant and I believe that this result can confirm that subject inversion structures are hard to process for children who develop Italian as a weak language.

5.6 Chapter summary and conclusion

In this chapter I have presented the results from the analysis of the longitudinal and experimental data from the four bilingual children and the control groups. The

initial analysis of the longitudinal data shows that there are few examples of postverbal subjects. Costanza is the child who produces the highest number of postverbal subjects also combined with different types of verbs. Her data also matches the monolingual Italian data in the distribution of preverbal and postverbal subjects. Overall, her subject production resembles that of the monolingual Italian children from GROUP A, who produced target-like answers.

Another result concerns the verbs used in the inversion structures. Consistent with Lorusso et al.'s (2004) results, the four bilingual children produce postverbal subjects mostly with unaccusative verbs. The only exception is found in Matelda's data, which shows an opposite trend (34.7% unaccusatives and 60.8% transitives). Overall, this analysis shows that the children who develop Italian as a weak language spontaneously produce very few postverbal subjects. However, the examples from the longitudinal data are too few to formulate a reliable hypothesis. I have suggested that the small number of samples in the data could be due to the lack of postverbal subjects in the input, the lack of overall input, cross-linguistic influence or the complexity of the structure. On the basis of the analysis of the spontaneous data, I have excluded the lack of postverbal subjects in the input and the cross-linguistic influence as possible causes. Due to the lack of sufficient samples in the data, I ran two experiments to elicit postverbal subjects with transitive and unaccusative verbs.

The first experiment tested the production of postverbal subjects with unaccusative verbs. The second experiment tested the children's ability to produce sentences containing not only subject inversion structures, but also reference to an antecedent. Both these structures require the activation of syntactic and pragmatic knowledge. The results of the two experiments confirm the initial prediction. Costanza, the child who develops Italian as a strong language, produces the highest number of postverbal subjects and performs better than the other children in both tasks. The children who develop Italian as a weak language have more difficulty in producing postverbal subjects in both tasks.

In order to further test the validity of this experiment in connection to the assumption on the weak language, the tasks were presented to control groups, including monolingual children and bilingual and monolingual adults. Adults and older monolingual children (GROUP A) produced the target answers, while younger monolingual children did not always produce full sentences. Another

interesting result is that children can produce VS order as early as 2 years of age. The second task was administered to the bilingual children again after 8-9 months, in order to observe possible changes. The results of this second test further confirm the existence of a relationship between the children's linguistic weakness and their ability to process structures at the interface between syntax and pragmatics. These results further demonstrate that the children who develop Italian as a weak language have difficulty producing both direct object pronouns and subject inversion. The reason could be the complexity of the structures or the interplay between the domain of syntax, which governs the word order and pragmatics, which governs the interpretation of the sentence and the introduction of information that is new to the hearer. The analysis of spontaneous and elicited data demonstrates that structures at the syntax-pragmatics interface can be analysed to test language dominance and could be further tested.

In this chapter, I have tested the hypothesis on the relationship between language dominance and the ability to process sentences containing structures governed by syntactic and pragmatic constraints. I have presented evidence from previous studies showing that subject inversion is a syntactic structure which is subject to pragmatic constraints, since it is used to express new information focus[60]. The evidence that adult L2 learners of Italian have difficulty producing VS structures (Belletti and Leonini 2004) suggests that these structures are hard to process because they require the activation of syntactic and pragmatic knowledge. On the basis of this evidence, which has been confirmed by other studies (Belletti et al. 2007, Bettoni et al. 2009), I have chosen to analyse the production of inversion structures. From this first analysis it emerges that postverbal subjects are less used than preverbal subjects, and they are used less by the bilingual children who develop Italian as a weak language than by those who develop it as a strong language. The longitudinal data was integrated with experimental data. The results of the experiments show that the children who develop Italian as a weak language have difficulty mastering the use of inversion structures. This evidence confirms the assumption that children who develop Italian as a weak language have difficulty mastering structures at the interface between syntax and pragmatics. The findings from the analysis presented in this chapter shed light on

[60] Also see Erteschik-Shir (2007) for an exhaustive account of the properties of information structures across several languages and López (2009) for their syntactic analysis.

the issue of language dominance in relation to processing resources in bilingual children.

CHAPTER 6

The role of the input in bilingual first language acquisition

6.1 Introduction

The analysis of the weak language presented in chapter 4 has shown that Matelda and Costanza develop Italian as a strong language, while Paolo and Francesca as a weak language. These results were further confirmed by the analysis of the production of subject inversion. Having established a method to assess weak language development, in this chapter I turn to exploring the causes of weakness, and examining the role of the input. The main hypothesis I wish to test is that both the quality and quantity of the input affect the acquisition of the minority language.

I assume that studying the input from a quantitative point of view is not sufficient to gain a full understanding of the relationship between input and dominance, because it only gives us an overview of the time the child spends being exposed to Italian. Therefore, it is necessary to explore the content of the parents' speech during the interaction with their children, in order to determine whether they actively engage in conversation and whether they provide a qualitatively rich input in terms of lexicon, syntax and morphology.

While in my analysis the input is considered the most significant linguistic factor that could justify a weak language development, it has to be taken into account that there may be other factors of a psychological and social nature which also affect 2L1 acquisition. However, in this thesis I only examine the role of the input, since it is more easily measurable. Moreover, it has been claimed by other researchers that the lack of input may cause linguistic imbalance in bilinguals (Schlyter 1993, Granfeldt and Schlyter 1994, Argyri and Sorace 2007).

So far, different methods have been employed to analyse the relationship between parents' input and children's linguistic development (see chapter 2). These are mostly based on questionnaires and on the analysis of the spontaneous interaction

between children and parents. The analysis that follows uses both methods to examine qualitative and quantitative aspects of the input. The quantitative analysis, based on the *Questionnaire on the linguistic background of the bilingual child* (Appendix B and C), aims to determine the amount of Italian time spent by the child with Italian speakers on a weekly and yearly basis, taking into account the child's daily routine and also the time spent in Italy every year. While the results of the questionnaire provide an overview of the time spent with Italian speakers, they do not explain the type of interaction taking place during this time. This gap is filled by the qualitative analysis, which is based on the parents' spontaneous data , and takes into account three factors, namely *output* (amount of words produced by the parent when interacting with the child), *vocabulary* (amount of word roots) and *syntactic complexity* (MLU in morphemes). These three factors have also been tested in Pancsofar and Vernon-Feagans (2006) study, which aimed at determining the differences between mothers' and fathers' input and their effect on the child's language development. However, my analysis differs in the type of data analysed and in the use of different samples (section *6.5.1*)

While previous studies have analysed the input by testing either spontaneous data or data from questionnaires, in my analysis I provide a comprehensive examination of the input by combining the two methodologies. The results of the quantitative and qualitative analysis are finally summarised in the Input Scale, which shows the total amount of qualitative and quantitative input the children are exposed to. By comparing the Weak Language Scale to the Input Scale it will be possible to determine whether there is a relationship between input and weak language development.

6.2 Italian in Ireland: external environment and home language use

According to the last census (CSO 2006), there are about 4 million people living in Ireland, of which 419,733 are non-Irish. The two official languages are Irish[61] and English. Irish is acquired as a first language only by part of the population, mainly residing in the Gaeltacht areas, but it is spoken as an L2 by a large part of the population. Ireland has seen a rapid economic growth since the 1990s, which resulted in considerable growth of the immigrant population. Today Ireland is a

[61] Since January 2007 Irish is also an official language of the European Union

multilingual country. The largest groups of non-English speaking immigrants are Polish (63.276), followed by Lithuanians, Latvians and Germans. The Italian community is among the 10 largest European communities living in Ireland (Rangone and Sgaggio 2007). With very few exceptions, the study of the Irish language is compulsory in primary and secondary education throughout the country. There are also schools where subjects are taught through Irish, but there are only a few bilingual schools that teach through languages other than English and Irish, and there is only one secondary school in Dublin where some subjects are taught through Italian. The Italian children involved in this research live in towns where Irish is not spoken and they are not exposed to any language other than English and Italian. They attend English-speaking pre-schools or day-care centres. Their families try to promote the use of Italian at home, since English is the dominant language in the children's environment. The four children come from middle class families in which the father works full-time and the mother part-time (to different extents), devoting the rest of the time to the care of their children. The parents try to promote bilingualism in the family and they generally stick to the *one-parent-one-language* strategy. In most cases, the Italian parent is bilingual and the Irish parent is monolingual. None of the English-speaking parents are fluent in Italian, and only one of them sometimes uses Italian with the partner. Overall, the children use English in a wider variety of contexts and with a higher number of speakers. The quality and quantity of input for each case study will be analysed in the following sections.

6.3 Methodology

Bearing in mind the characteristics of the external environment, I will now look more closely at the input focusing on each case-study. In sections 6.4 to *6.4.4*, I present the quantitative analysis, based on the results obtained from the *Questionnaire on the linguistic background of the bilingual child*, (see Appendix B and C). The questionnaire is compiled by the parents of the four bilingual children (see Chapter 3), who are asked to observe their family's linguistic behaviour during a normal week. On the basis of the questionnaire, I aim to estimate the average quantity of exposure to Italian.

6.4 Exposure to the two languages: quantitative analysis

The results presented in the following sections are based only on the answers given in the questionnaire. To keep the anonymity of the children and their families, all names have been changed and no information on their identity is provided.

6.4.1 Quantity of input: Paolo

Paolo was 3;1 at the time of the first recording. He has an Italian mother, an Irish father and no siblings. His parents follow the *one-parent-one-language* strategy and they speak only English to each other. They were initially concerned because Paolo was not producing full sentences in Italian and he was constantly replying in English when addressed in Italian. His Italian was, in their opinion, very limited. Paolo has always lived in Ireland, where he has attended a playschool since he was 12 months old for 4 days a week for a total of 36 hours a week. He spends about 15-20 hours a week with his mother and 10-15 with his father. He is exposed to an average of 2-3 hours of Italian and 10-13 of English everyday. He speaks mostly English and sometimes he uses it when speaking to his mother. He uses Italian daily with his mother and sometimes on the phone to the family in Italy. He also spends at least one month in Italy every year. The father is not fluent in Italian. Both parents want Paolo to be fluent in Italian and they try to promote the use of the language in the home by reading stories, watching cartoons and talking. The following figures show a graphic representation of the quantitative analysis of the input, based on the parent's responses to the questionnaire.

Figure 6.1 Average number of days spent in Italy and Ireland each year - Paolo

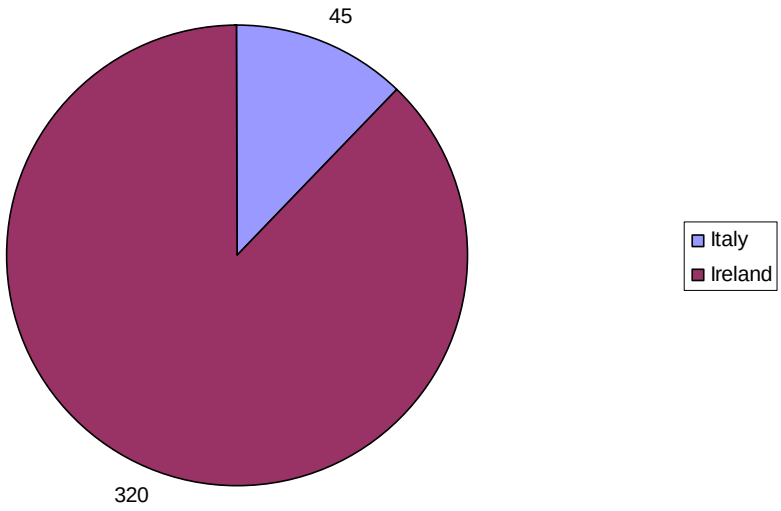


Figure 6.2 Average percentage of weekly exposure to each language in Ireland - Paolo

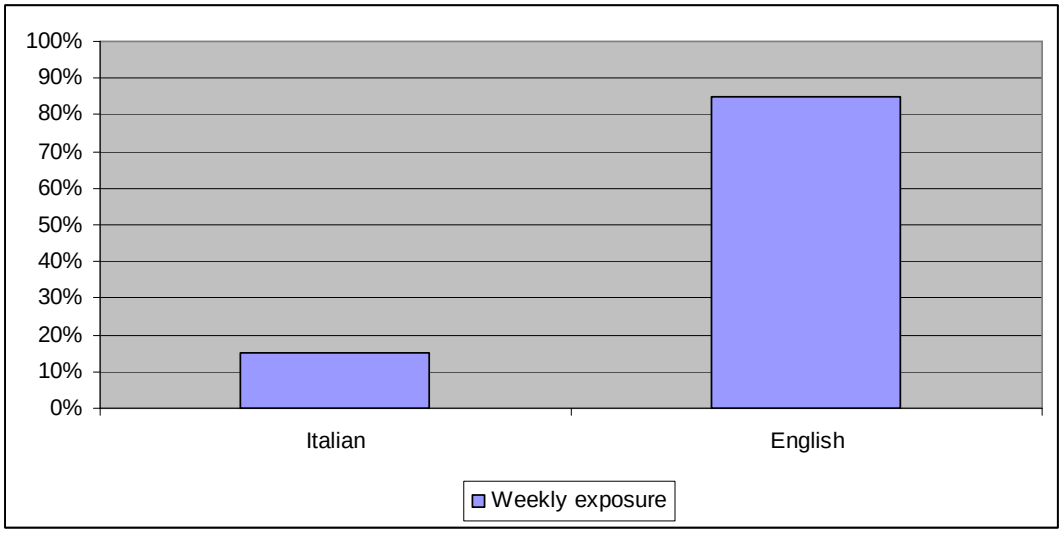
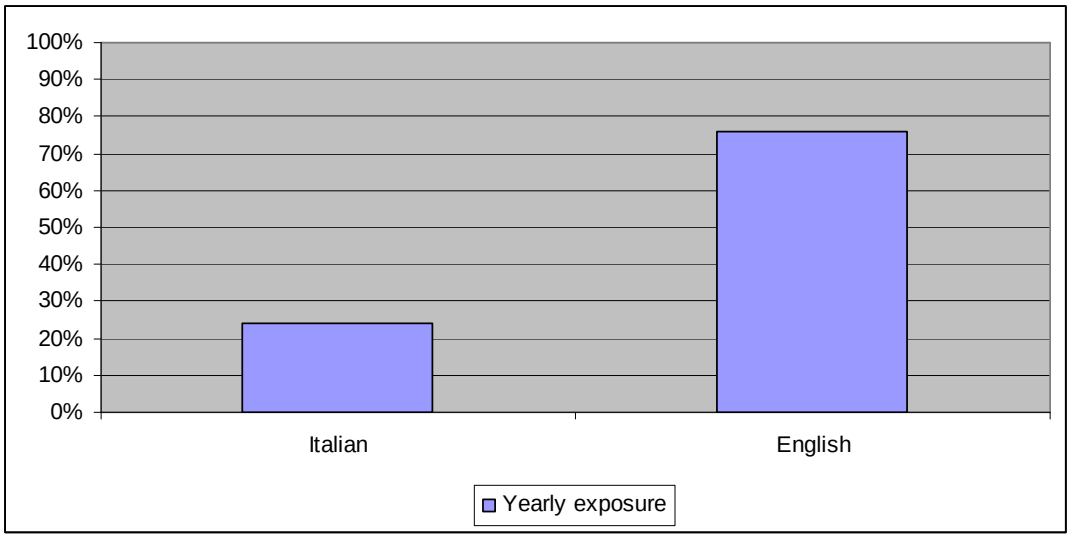


Figure 6.3 Average percentage of exposure to each language in one year - Paolo



6.4.2 Quantity of input: Francesca

Francesca was 2;4 at the time of the first recording. She has an Italian mother, an Irish father and no siblings. Her parents follow the *one-parent-one-language* strategy. They speak mainly English to each other, but sometimes they try to speak Italian. Francesca has always lived in Ireland, where she has been attending a crèche since she was 9 months old for 5 days a week for a total of 25 hours a week. She spends about 25 hours a week with her mother and 20 with her father. She uses English in the crèche, with her father and his family and Italian with her mother and her family. She has also spent at least one month in Italy every year since her birth. Her father can speak Italian, but he is not fluent. Both parents try to promote the use of Italian and want the child to master the language.

Figure 6.4 Average number of days spent in Italy and Ireland each year - Francesca

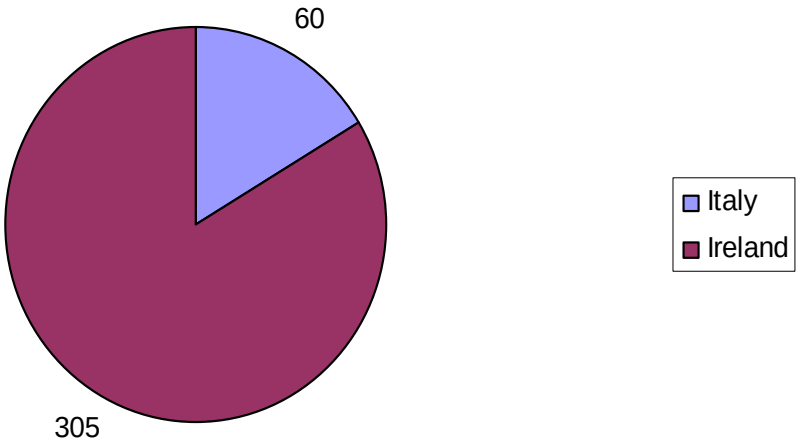


Figure 6.5 Average percentage of weekly exposure to each language in Ireland - Francesca

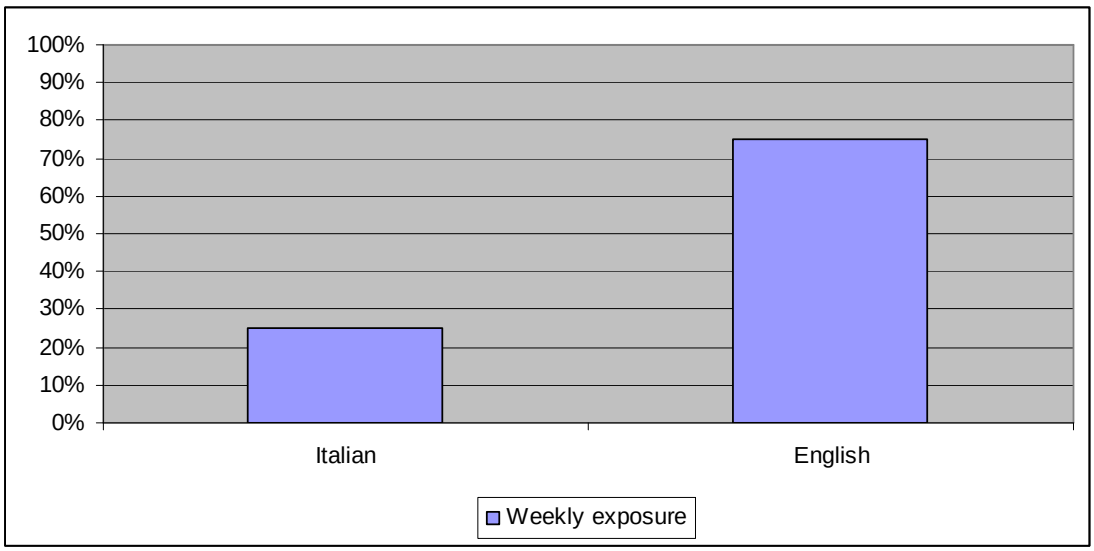
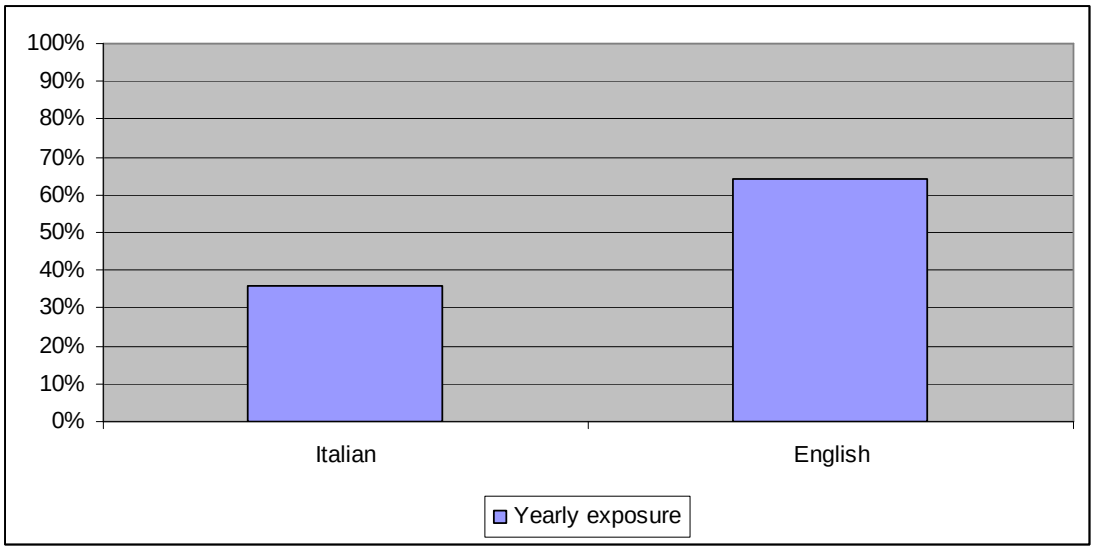


Figure 6.6 Average percentage of exposure to each language in one year – Francesca



6.4.3 Quantity of input: Costanza

Costanza was 1;11 at the time of the first recording. She has an Italian mother, an Irish father and no siblings. Her parents follow the *one-parent-one-language* strategy and they speak only English to each other. Costanza has always lived in Ireland, where she attended a day-care centre since she was 6 months old for 2 days a week for a total of 16 hours a week. When the parents work, she is taken care of by other family members who only speak English. She spends about 25 hours a week with her mother and 8 with her father. The child is exposed to about 4 hours of Italian and 8 of English every day. She uses English in the créche, with her father and his family and Italian with her mother and her family. She spends 1 or 2 months in Italy during the summer months. The mother promotes the use of Italian by speaking only Italian to Costanza and by using Italian games, stories and movies. Her mother has made every effort to provide as much input as possible in the daily routine and has also tried to spend time with the Italian family whenever possible.

Figure 6.7 Average number of days spent in Italy and Ireland each year - Costanza

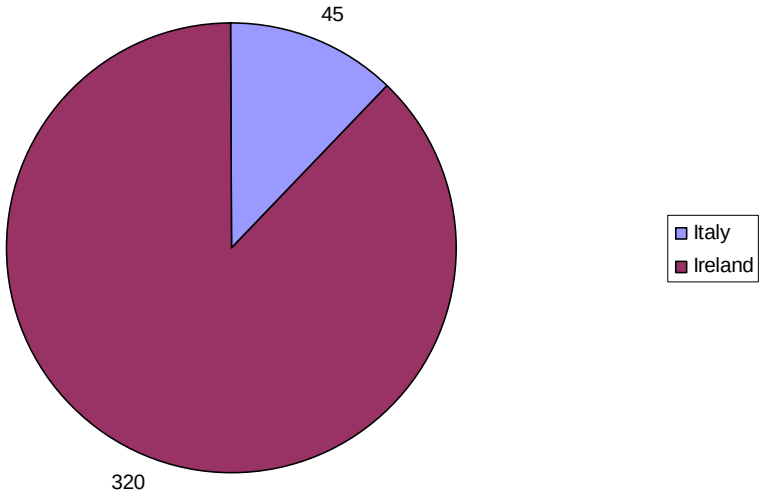


Figure 6.8 Average percentage of weekly exposure to each language in Ireland - Costanza

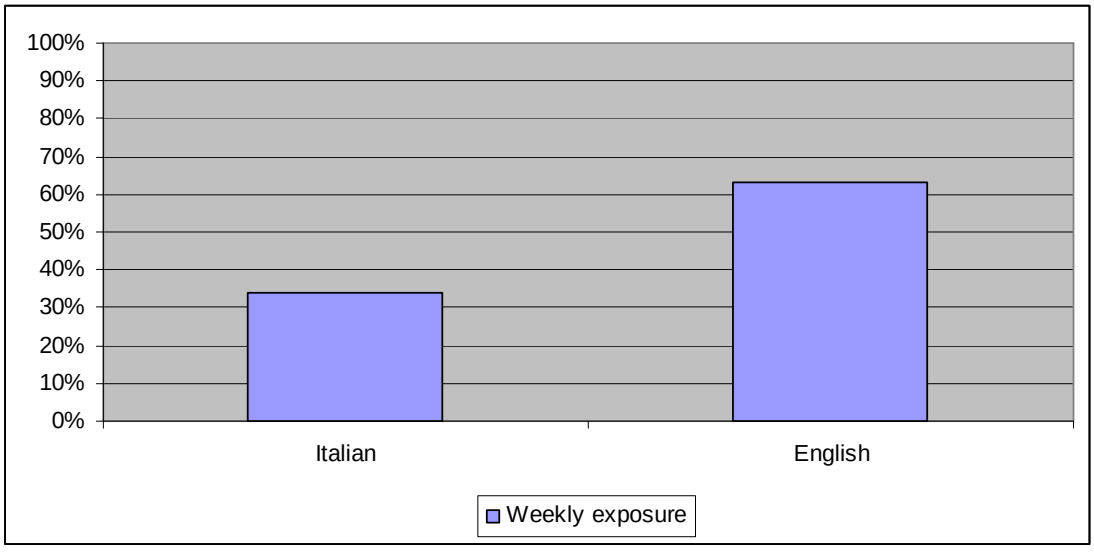
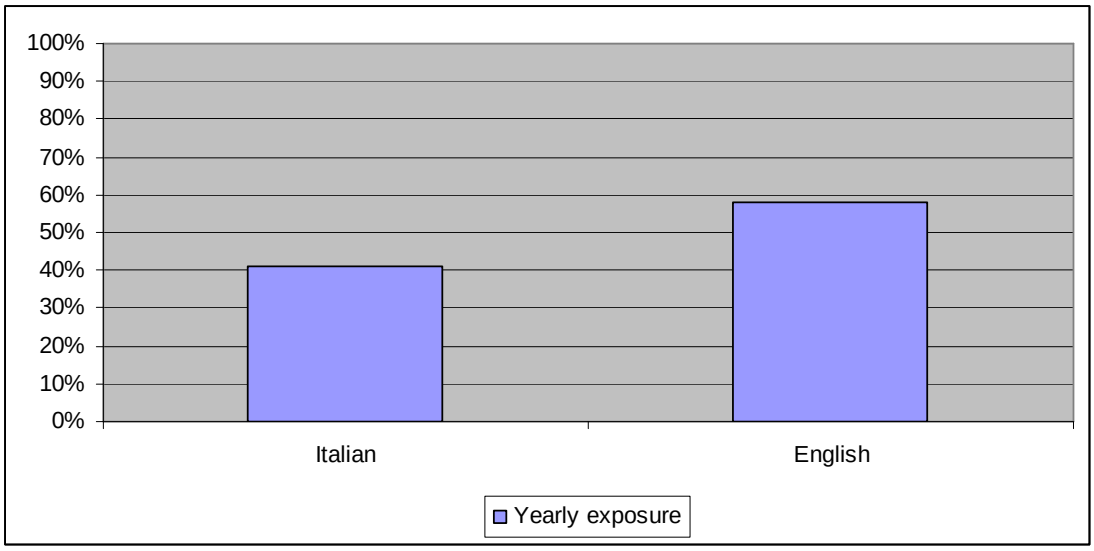


Figure 6.9 Average percentage of exposure to each language in one year - Costanza



6.4.4 Quantity of input: Matelda

Matelda was 2;6 at the time of the first recording. Both her parents are Italian and they moved to Ireland as adults. She has a younger brother who is a toddler and is not yet able to talk. When addressing the child, the father only speaks Italian, the mother also uses mixed utterances. The parents always speak Italian to each other. Matelda always lived in Ireland, and she did not attend any day-care. She was minded for a few hours every day by an Italian au-pair. Her mother brings her to community playgroups to expose her to English. Matelda spends about 80 hours a week with her mother and 40 with her father. She is exposed to 10 hours of Italian and 4 of English every day. She uses Italian at home and English with the children in the neighbourhood. Matelda also spent an average of three months in Italy every year since her birth.

Figure 6.10 Average number of days spent in Italy and Ireland each year - Matelda

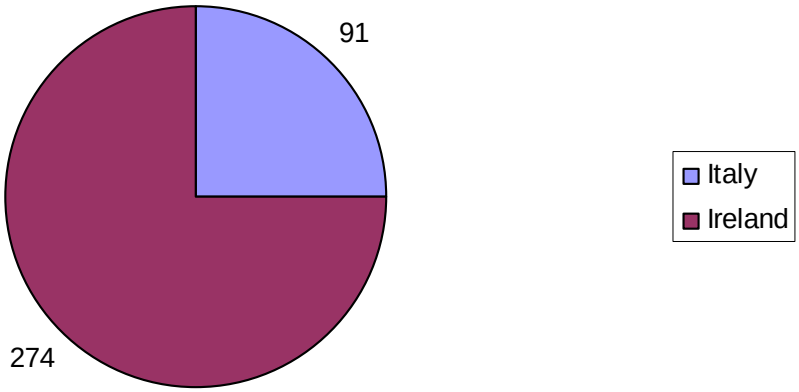


Figure 6.11 Average weekly exposure to each language in Ireland - Matelda

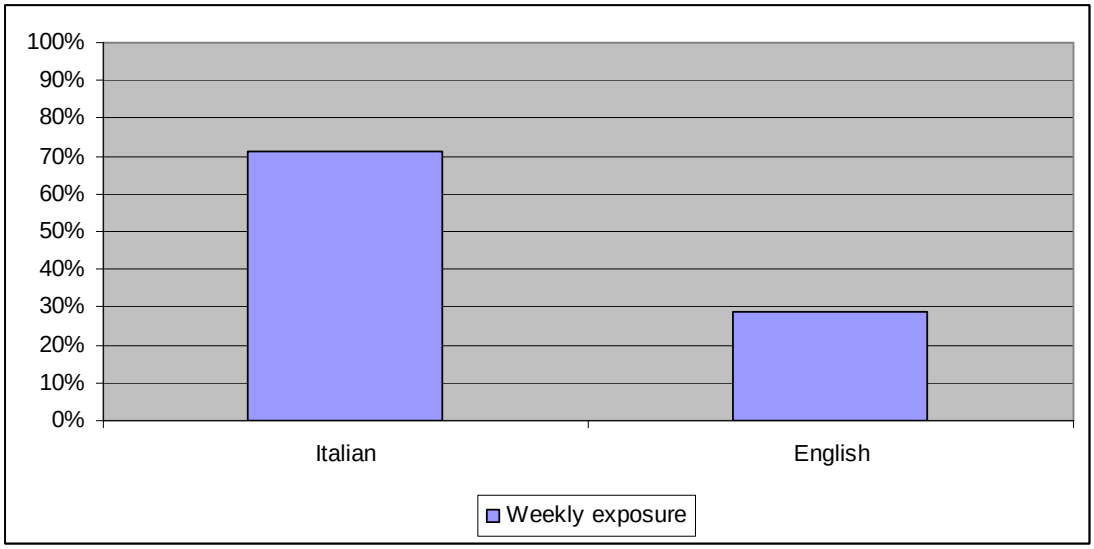


Figure 6.12 Average percentage of exposure to each language in one year - Matelda

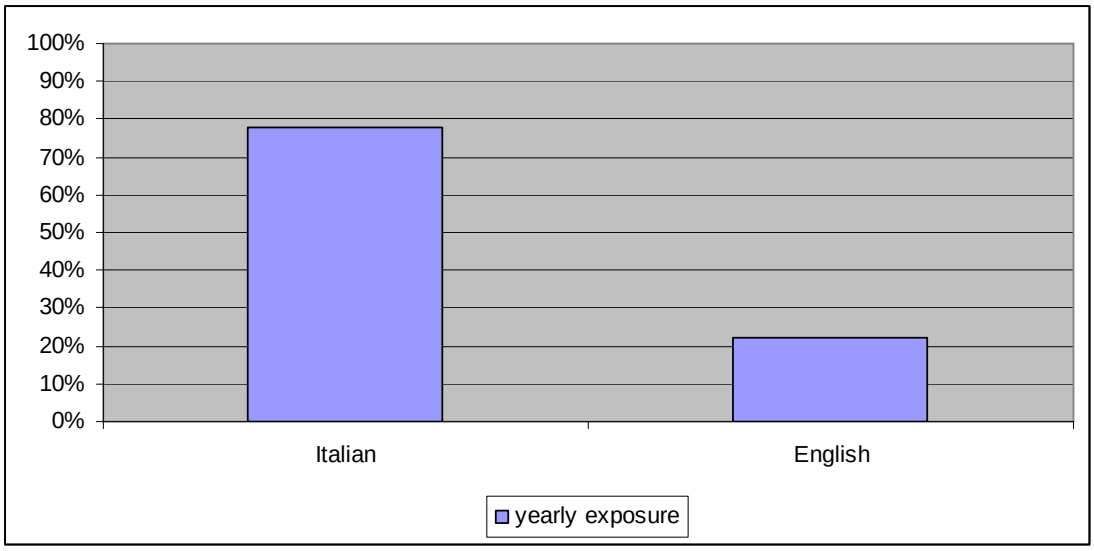


Figure 6.13 and 6.14 provide a general overview of the data across the four children. The results of the quantitative analysis show that the children involved in this study are exposed to Italian less than 40% of the time, and only Matelda is exposed to Italian more than to English.

Figure 6.13 Average exposure to English and Italian on an average week

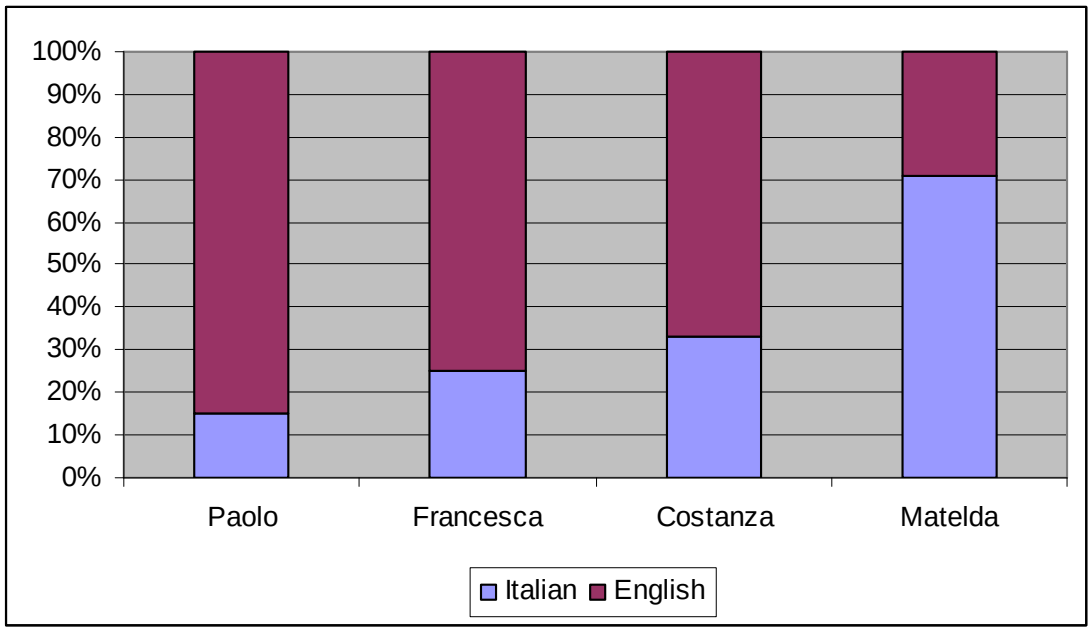
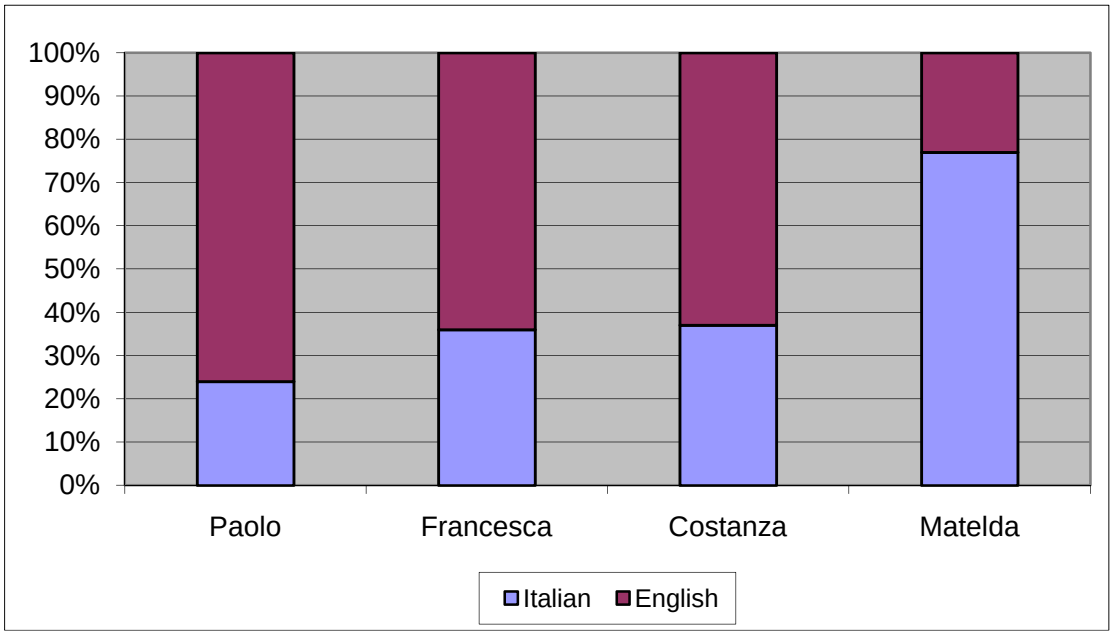


Figure 6.14 Average exposure to English and Italian in a year



6.5 Exposure to the two languages – Quality of input

As I have shown in table 2.5 (chapter 2), different methods have been used to analyse the quantity and quality of the input. While the quantitative analysis measures the amount of hours or days of exposure to the language, the qualitative analysis explores the actual content of the input. The factors that are generally taken into account in qualitative studies are the mean length of utterance (MLU), the syntactic complexity, the lexical variety and also other extra-linguistic aspects. The quality of the input is generally analysed on the basis of samples of conversation between the children and their parents. A factor that is often taken into account in studies on child directed speech (CDS) is the socio-economic status of the parents (Pancsofar and Vernon-Feagans 2006; Huttenlocher et al. 2002; Windsor et al. 2007). However, this aspect will be excluded in the present study since it does not constitute a significant factor[62.]

In order to analyse the quality of input it is necessary to consider what elements make some parents' speech "richer" than others. The factors that will be taken into consideration to measure the quality of the input are output, vocabulary and complexity of utterances. Previous research suggests that the combination of

[62] All the children involved in the present research come from middle class families; all parents have a university degree. In each family, the father has a full-time job and the mother has a part-time job, mostly working 4 days per week.

these qualitative factors contributes to the development of early language skills (Bornstein, Haynes and Painter, 1998; Hart and Risley, 1995, Hoff-Ginsberg, 1991). *Output* is the number of utterances produced by the parents when addressing the child, and it is calculated by counting the total number of verbal utterances. Another important factor to be taken into account is the amount of *vocabulary* produced by the parents, which is calculated by counting word roots, and the complexity of their utterances, represented by the average MLU calculated counting the number of total morphemes. *Context of use* refers to the situation in which the language is spoken. According to this analysis, a qualitatively rich input is provided by parents who produce morphologically and syntactically complex sentences, use a varied vocabulary and engage in conversation as much as possible while with the child.

In the analysis that follows, the names of the parents are not displayed. Since the following chapters will focus on the children's data, I consider the use of the child's name more appropriate in order to make the data more easily comparable at a later stage. I will therefore use the label *Mother of* followed by the name of the child.

6.5.1 The samples

In the analysis that follows, I examine three factors, namely output, vocabulary and syntactic complexity. To analyse each factor taking into account possible changes through time, I select three samples, one from the beginning, one from the middle and one from the end of the recording period. It has to be also taken into account that the analysis of output, vocabulary and syntactic complexity are not based on the same sample, because each analysis requires a different sampling methodology. The output therefore is calculated on the basis of 10 minutes of continuous interaction between the parent and the child; the vocabulary and the MLU are calculated on the basis of 100 utterances.

6.5.2 Output

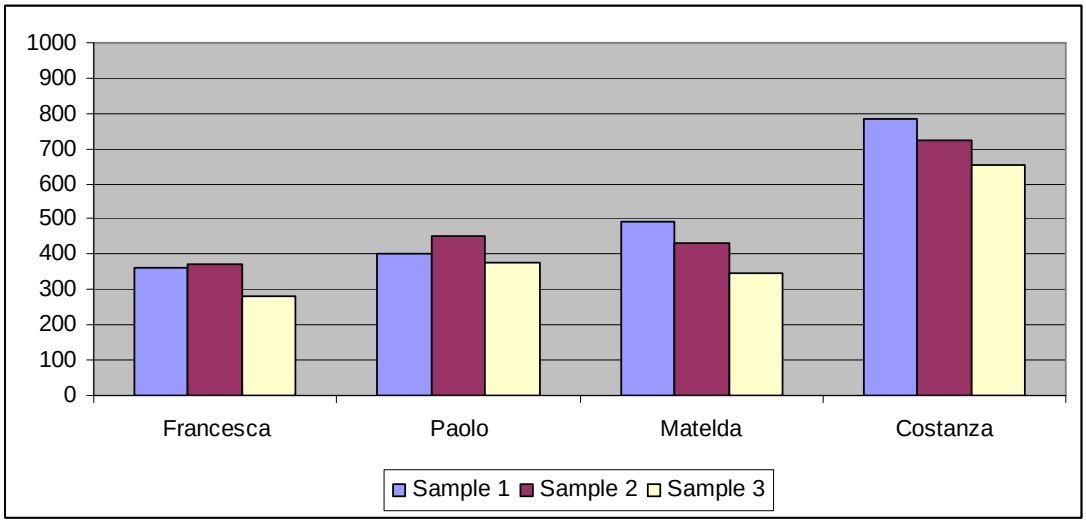
The output is analysed in order to determine how much linguistic input is provided during the interaction between the parent and the child. This is an important qualitative factor, since it can show whether parents engage in conversation and whether they spend much time speaking to the child.

The output is calculated by adding up the number of utterances produced by the parent over a period of 10 minutes during 3 free play sessions, displayed in chronological order (Table 6.1 and Figure 6.18). The choice of samples was based on the following criteria: it was not at the beginning of the recording, when the parents are more aware of the presence of a recorder, it did not involve the reading of a story, and it was an interaction as spontaneous as possible. I chose to analyse only 10 minutes from the total recording because I found that it was possible to collect samples of this length from all stages[63].

Table 6.1 Output

Mother of	Number of words uttered in 10 min. Sample 1	Number of words uttered in 10 min. Sample 2	Number of words uttered in 10 min. Sample 3	Average
Francesca	364	370	280	338
Paolo	400	450	379	409.6
Matelda	490	430	346	439.6
Costanza	786	725	653	721.3

Figure 6.15 Output



The results show that when engaging in conversation or play with their children, the parents produce between about 300 and 700 words in each 10 minute sample.

[63] Pancsofar and Vernon-Feagans (2006) analyse this factor on the basis of 20 minute samples.

Costanza consistently receives almost twice as much input as the other children. It is interesting to note that there seems to be an inverse proportion between the number of words produced by the parents and the children's age. As the children's MLU increases, there are more utterances produced by the child and fewer by the parent. The results of the output show differences among parents and they constitute a first significant step in exploring the quality of the input.

6.5.3 Vocabulary

The second factor considered in this analysis is the amount of vocabulary in the input[64]. According to some researchers, lexical learning occurs mostly when there is mutual engagement in conversation (Hoff and Naigles 2002) and input frequency has an impact on the child's acquisition of the lexicon (Smith 1999). Also the production of a high number of word types in the input has been found to positively influence the child's lexical comprehension and production (Bornstein, Haines and Painter 1998). Hoff and Naigles (2002) argue that children's vocabulary development is influenced by 'sheer frequency of presentation, number of different words, and richness and variety of linguistic environments in which the words are placed' (p. 423).

I calculate lexical variety on the basis of the number of different word roots in the conversation between the parent and the child in three different samples. The analysis of the lexical variety is based on 100 utterances from three samples selected on the basis of the criteria enumerated in the previous section. Unintelligible words and fillers are omitted. The results are based on the number of different word roots in each sample (Table 6.2).

[64] This factor is also included in Pancsofar and Vernon-Feagans' (2006) analysis. However, they do not take into account samples from different stages.

Table 6.2 Vocabulary

Mother of	Sample 1	Sample 2	Sample 3	Average
Francesca	190	226	182	199.3
Paolo	286	202	228	238.6
Matelda	184	154	226	188
Costanza	272	280	260	270.6

6.5.4 Syntactic Complexity

The method that is commonly employed for assessing the complexity of utterances is the calculation of the mean length of utterance (MLU) in words and morphemes (Brown 1973). While the analysis of the children's MLU presented in chapter 3 was based on the number of words (in order to make the results comparable to other studies), the parents' MLU is calculated on the basis of the number of total morphemes. This type of analysis reflects the syntactic complexity, since it accounts for the number of total morphemes. Adverbs and uninflected forms are counted as a single unit, and inflected forms are counted as two morphemes, since they contain a morpheme that conveys morphological marking of singular/plural and masculine/feminine. The MLU analysis shown in table 6.3 is based on 100 utterances taken from 4 different samples.

Table 6.3 Parents' MLU

Mother of	Sample 1	Sample 2	Sample 3	Average
Francesca	4.52	6.4	7.64	6.18
Paolo	7	7	7.5	7.16
Matelda	8.52	5.6	7.4	7.17
Costanza	8.92	8.92	8.94	8.92

The results on vocabulary and MLU show that there is a certain degree of variation among the parents. Overall, Costanza is the child who receives on average the highest quality of input. In the following section, I will present the results on the basis of a scale created for comparing the children.

6.6 Context of use and attitude

The factors analysed so far are numerically quantifiable. However, there are many other factors that constitute part of the input, which are not easily measurable. One of these factors is the context of use. Bilingualism research has shown that children develop pragmatic competence that enables them to choose the language to use in a specific context from the earliest stages (Ritchie and Bhatia 2004: 339). The analysis of the context aims to provide an insight into the use of language by the child. The children taking part in this study are brought up speaking English and Italian, but since they are raised in an English speaking country, they are likely to be exposed to English in a wider variety of contexts and situations. Overall, on a daily basis the children are mostly exposed to adult language, mainly at home or in one-to-one situations rather than in social contexts. From the questionnaire, it emerges that the children are exposed to Italian mainly at home with the parents and occasionally with Italian people, who are mostly adults (extended family, friends, childminders or visitors). They are also exposed to Italian to a variable extent during their trips to Italy.

Table 6.4 Answers to the question: “How do you promote the use of Italian in your family?”

Mother of	Speak Italian	Read Books	Watch TV	Trips to Italy	Meet other Italians	Sing Songs
Francesca	✔	✔			✔	✔
Paolo	✔	✔	✔			
Matelda	✔	✔	✔			
Costanza	✔	✔	✔	✔	✔	

Table 6.4 represents the answers given by the parents. The questionnaire did not present a multiple choice, and these are the answers spontaneously provided by

the parents. These responses can be interpreted as an indication of the motivation the parents have to promote the use of Italian. Children realise from very early on that there is one language spoken by the majority of the people in the country where they are raised, and there is a minority language that is relegated to limited contexts of use and to a limited number of speakers in their environment. It is difficult to determine whether a child enjoys speaking one language more than the other or whether motivation can be reliably measured. Even though there is no evidence showing the role of motivational factors in 2L1 acquisition, there is a possibility that the child will not develop interest or attachment towards one of the two languages. This behaviour could be developed independently from any external factors, or it could be a reflection of the family's attitude. While L2 studies can employ questionnaires or tests to analyse emotional factors, studies on young bilingual children have to rely on the parents' judgement or on observation of the child's behaviour. On the basis of these judgements, the children analysed in this thesis seem to have developed a positive attitude towards Italian, they associate it with family, holidays, play and fun and they often talk about their experiences in Italy and their time spent with their families.

The older the children get, the easier it is for them to explicitly communicate their feelings towards the language. However, parents themselves can communicate their own feelings towards the language and culture by making choices in their daily language use and by selecting linguistic strategies. All the parents taking part in this research have adopted the *one-parent-one-language* strategy and hardly ever mix the two languages when talking to their child. Moreover, the results of the questionnaire showed that both the Italian and the Irish parent have a positive attitude towards the minority language and they use different strategies to promote its use in the family. As shown in Table 6.5, the two strategies that are used by all parents are speaking and reading books. Many other social and emotional factors could be taken into account. Some of these are the attachment to the Italian family, the relationship with other speakers of Italian, the contact with other children, and many more. These factors of a more social and psychological nature are not easily quantifiable, and are not included in the analysis that follows.

6.7 Assigning values

By adding the values assigned to quantitative and qualitative factors, it is possible to obtain a result that indicates the total amount of qualitative and quantitative input received. This result will later be used to determine whether there is a relationship between input and dominance[65]. By combining all the factors analysed so far, I aim to find out whether a low amount of input (quantitative and qualitative) results in weak linguistic development.

In order to be able to add up the factors, I use the same criterion applied in the analysis of the weak language. The values are assigned independently for each factor by examining the data and determining minimum and maximum values, which would correspond to a scale from 1 to 10 (see Appendix D). For example, the quantity of input is calculated by averaging the results of the amount of input received over a week in Ireland and the overall yearly result, which would also take into account the time spent in Italy. A child could be exposed to 90 hours of language (this result emerges from the questionnaire), of which 20 is Italian and 70 is English in a normal week in Ireland. The same child could be visiting Italy every year for one month, and during that month be exposed to 80 hours of Italian and 10 of English a week. Therefore, the *quantity of input* factor takes into consideration the amount of Italian input considering both the exposure on an average week and the time spent in Italy. On the basis of the data from the questionnaire, I established that the maximum value assigned to the quantity of input would be 100. Obviously, a result between 90 and 100 would mean that the child is mostly exposed to one language only. While this does not apply to the children under examination in this thesis, it is a possible result that might emerge in other studies.

The prediction is that if the factors taken into account are relevant to determine balance, the lower the score, the more likely the child is to develop Italian as the weak language. Not many studies have so far presented a comprehensive analysis involving quantitative and qualitative factors, therefore there is not yet a valid method of assessment of the input that could be compared to that used in the

[65] A limitation of this method lies in the possibility of determining whether some factors are more significant than others, due to the small amount of data.

present study. Moreover, as in the analysis of the weak language, more data is needed to confirm the validity of the methodology.

Table 6.5 Results of the sum of values[66]

Factors	Costanza	Paolo	Matelda	Francesca
Quantity of Input	3.5	1.5	7.5	3
Output	7	5	4	3
Vocabulary	5.5	5	4	4
Average MLU	5.5	4	4	3
TOTAL	21.5	15.5	19.5	13

Since there are four factors accounted for, the minimum final result is 4 and the maximum is 40. I assume that children who get a score between 18 and 22 receive sufficient input to develop Italian as a strong language (fig. 6.19). This assumption is based on the assessment of the bilingual data examined in this thesis, but more data is necessary to further confirm the validity of this hypothesis.

On the basis of the factors taken into account and the values assigned to them, Costanza and Matelda receive a higher total quantity of input than Paolo and Francesca. It has to be noted that these results are based only on 1 year in the child's life and that the amount of input in the two languages can change. This model therefore reflects the performance in a selected time-frame and it is not expected to make predictions on the children's linguistic development beyond the period analysed. However, I believe that any future change in the child's developmental trend will continue to be strongly influenced by the input.

6.8 Input and weak language development

The initial hypothesis formulated in this thesis is that the development of the minority language is affected by the input. In order to test this hypothesis, I have presented two assessment methods, one that calculates the sum of the quality and

[66] The name of the children is used instead of the parents'. In this case it should be read as "the mother of".

quantity of the parents' input, and one that assesses the children's weak or strong development of Italian. The following figures show the results of the two analyses.

Figure 6.16 Input Scale

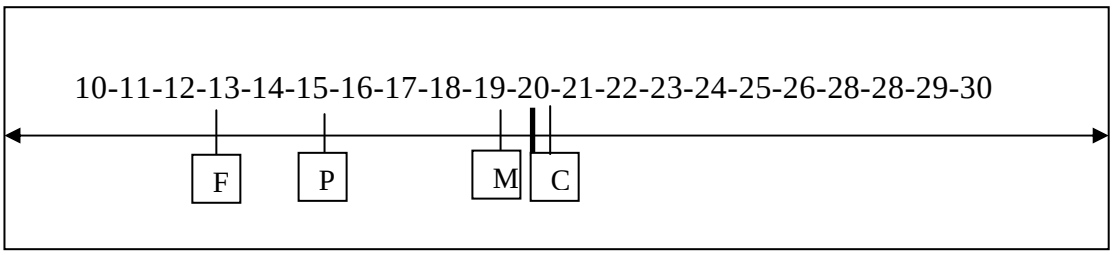
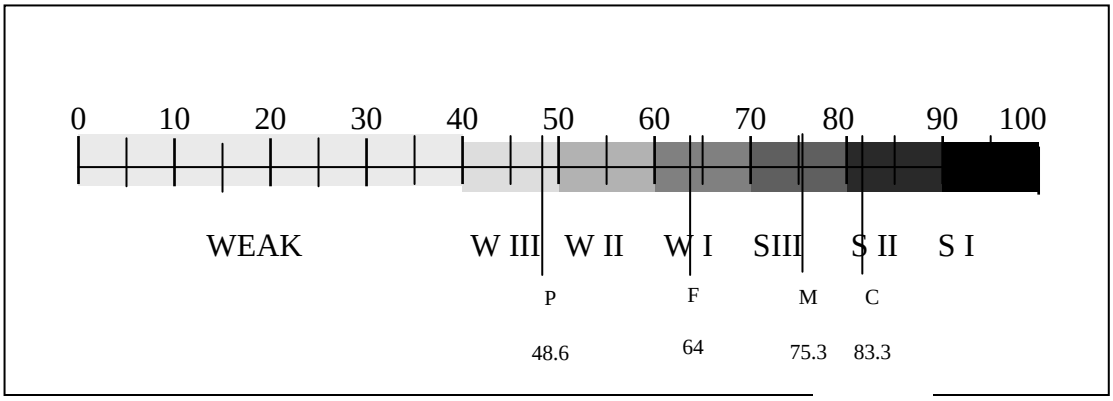


Figure 6.17 Weak Language Scale



If we look at the two figures above, we can see that the initial prediction is partly validated. The results of the assessment of the input suggest that Costanza and Matelda are exposed to a high amount of input and they develop Italian as a strong language.

Since I am dealing with a relatively small population, it is possible to analyse each individual case, in order to determine whether there is a correlation between a quality/quality and weak language development. This analysis is based on the correlation of pairs of factors from the data on the input and those on the output. If we compare Francesca and Costanza, who received a similar quantity of input, we can see that the qualitative values of Costanza's input are higher. The data also shows that there is a correlation between the quantity of input received and the rate of acquisition of determiners. Moreover, Costanza and Francesca are exposed to the same (or similar) quantity of input and develop at different rates. On the other hand,

the cases of Paolo and Matelda show that the quantity of the input is a significant factor. Matelda is exposed to a considerably higher quantity of input compared to Paolo. Even though Paolo receives a qualitatively rich input, his development is weak, he produces a high number of target-deviant forms and his MLU is low compared to other age-matched children.

It emerges from the analysis of the input and that of the weak language that both qualitative and quantitative aspects are relevant for the acquisition of the minority language, but it is not possible to determine their statistical significance of each factor due to the small amount of data. However, the results from the four bilingual children and their parents show that both quantity and quality should be analysed, and their interplay provides a rich input for the child.

The comparison between the Input Scale and the Weak Language Scale shows that Costanza and Matelda, who received the most consistent input, develop Italian as a strong language and their level on the scales is almost proportional. However, the same does not apply to Paolo and Francesca's data (Paolo receives more input, but his development is weaker than Francesca's). Different hypotheses can account for this result. The first may be that a quantity of input that goes below the 20% of the total exposure to language may be too low to develop a strong language (Paolo's exposure to Italian is lower than 20% – see figure 6.13). It is also possible to hypothesise that even if the quality of the input is high, the very low amount of time of exposure to a language may result in its development as a weak language. To test this hypothesis, it would be necessary to analyse data from a child who receives the same amount of input. The second hypothesis is that more or different factors may have to be taken into account to find a relationship between the input and the weak language. In this thesis I have only considered the input as a factor affecting the development of the weak language. In addition, it is possible that other linguistic and non-linguistic factors may have to be taken into account. This and other limitations of this analysis will be discussed in the final chapter.

Even though the results of the two models presented to analyse the input and the weak language do not perfectly match, it is important to point out that the overall result achieved confirms the initial prediction, demonstrating that there is a relationship between the total input and the development of the minority language as weak or strong.

6.9 Chapter summary and conclusion

Starting from the assumption that parental input plays a major role in the child's acquisition of the minority language, in this chapter I have explored both the quantitative and qualitative nature of the input, basing the analysis on naturalistic data and on the results of a questionnaire. The questionnaire provided information that was used to determine the quantity of input each child received. It also included questions relating to the attitude of the parents and their efforts to promote the Italian language in the family. However, these factors of social and environmental nature were not included in the analysis, since they are not easily quantifiable.

Adapting the methodology used by Pancsofar and Vernon-Feagans (2006), I analysed the role of the input on the basis of three qualitative factors, namely output, vocabulary and complexity of parents' utterances.

The sum of qualitative and quantitative factors was used to build the Input Scale (figure 6.19). This scale treats each factor as equally significant, and the results are based on the sum of the values given to each variable on the basis of a scale. This model was created on the basis of Italian data and it may be applied to other languages, as long as the values of the factors taken into account (such as MLU values) are modelled on the language under examination. The scale was constructed in order to present data on the input, which can be compared to the data on the children's linguistic development presented in chapter 4. The scale shows the total quantity and quality of input each child receives. The results show that Francesca and Paolo receive an amount of input lower than Costanza's and Matelda's.

This finding confirms the initial hypothesis that the input has an effect on the child's linguistic development, and that insufficient input might result in development a weak language. The method applied to analyse the input is useful for comparing the parent's with the child's data and it is reliable especially in contexts where one parent is the main source of input. Since there is still not a full account of the causes of weak language development, this analysis could serve as a starting point for further research.

CHAPTER 7
Discussion and conclusions

7.1 Introduction

In this final chapter, I present the findings of the thesis and an overview of the methodological and theoretical implications of this research, followed by an evaluation of the contribution to existing knowledge and recommendations for future research.

7.2 Review of chapters

In this section I provide an outline of each of the chapters presented in this thesis.

Chapter 1: Introduction and theoretical assumptions

In this chapter I set out the aims of the thesis by introducing the main research questions, the data, the methodology and the theoretical background.

Chapter 2: The study of the weak language

In this chapter, I evaluated the methods and the results of previous studies on the weak language. I examined some of the most relevant findings that have emerged in the literature on the weak language, considering the methodologies employed and the factors analysed. I also presented research findings that show that some bilingual children and adults might exhibit difficulty processing structures at the interface between syntax and pragmatics. I hypothesised that the children who develop Italian as a weak language might have difficulty producing postverbal subjects. Finally I evaluated the results of studies on the role of the input in L1 and 2L1 acquisition, suggesting that there might be a relationship between the input and the development of a weak or strong language.

Chapter 3: Overview of the longitudinal data

In this chapter, I introduced the linguistic background of each bilingual child participating in the longitudinal study. After explaining the methodology of data collection, I examined each case-study, focusing on the characteristics of the

linguistic environment and the amount of exposure to Italian, providing examples of interaction between the children and the Italian interlocutors at different stages of development.

Chapter 4: Strong and weak development of Italian

In this chapter, I presented the methodology used to determine whether Italian is a weak or strong language. The factors included in the analysis were rate of acquisition, code-mixing, MLU, vocabulary, use of target-deviant forms and subject distribution. These factors were chosen because they cover different areas of linguistic competence and they can provide a comprehensive and longitudinal view of the child's linguistic development. Another important methodological choice was the selection of comparative data. In this chapter, I used the data and results from analyses on monolingual and bilingual children, in order to present a wider perspective on the development of Italian in different groups of speakers. The results showed that there is variation among the four bilingual children, and that the factors analysed can be further tested for the assessment of language dominance.

Chapter 5: The weak language and the syntax-pragmatics interface

In this chapter, I confirmed the initial hypothesis that bilingual children who develop Italian as a weak language have difficulty mastering the use of postverbal subjects. The initial analysis of the longitudinal data showed that there are few examples of postverbal subjects, and that the children who develop Italian as a weaker language spontaneously produce fewer than those who develop Italian as a strong language. Due to the lack of sufficient samples in the data, I ran two experiments to elicit postverbal subjects with transitive and unaccusative verbs. The results of the two experiments confirm the initial prediction. Costanza, the child who develops Italian as a strong language, also produced the highest number of postverbal subjects and performed better than the other children in both tasks. The children who developed Italian as a weak language had more difficulty in producing postverbal subjects in both tasks. These results confirm the hypothesis that children who develop Italian as a weak language have difficulty mastering structures at the interface between syntax and pragmatics.

Chapter 6: The role of the input in bilingual first language acquisition

In this chapter I explored both the quantitative and qualitative nature of the input, basing the analysis on naturalistic data and on the results of a questionnaire. Adapting the methodology used by Pancsofar and Vernon-Feagans (2006), I

analysed the role of the input on the basis of three qualitative factors, namely output, vocabulary and complexity of parents' utterances. In addition, I analysed the results from a questionnaire on the child's linguistic background, which provided data on the quantity of input. The results of this analysis showed that the children who are exposed to the largest amount of quantitative and qualitative input develop Italian as a strong language.

7.3 Concluding remarks

This thesis provides a new insight into the bilingual language faculty. Following the generative framework, I assumed that there is a universal basis underlying the human language making capacity. In addition, the external environment provides the child with evidence to set the language-specific parameters and to acquire the properties of the language. One of the three main research questions addressed in this thesis concerns the role of this evidence in bilingual first language acquisition. On the basis of previous research findings, I assumed that a child who is not exposed to a sufficient amount of qualitative and quantitative input will develop one language as a weak language. As I have shown in chapter 6, children who are raised in bilingual families in a predominantly monolingual country are likely to be exposed to the two languages to a different extent both in quantitative and qualitative terms. The results from the analysis of the input in the four case studies presented in this thesis show significant differences in the exposure to the Italian input. Even though there is not yet any indication in the literature as to how much input bilingual children need to develop the two languages equally (or almost equally), it emerges from the analysis presented in chapter 6 that both quality and quantity of input are necessary. Due to the limited amount of data, it was not possible to establish the minimum amount of quantitative and qualitative input a child requires to develop Italian as a strong language. However, this method can constitute a starting point for larger-scale studies on the relationship between input and language development in bilinguals.

Another important question addressed in this thesis concerns the relationship of the input and the weak language. Having analysed different aspects of the children's linguistic development, I compared the results from the Input Scale to those of the Weak Language Scale. This comparison showed that the children who received the largest amount of qualitative and quantitative input develop Italian as a strong language. The role of the Weak Language Scale is not only to compare the children's

results, but also to show that there are different levels of linguistic weakness and strength. As I suggested before in this thesis, it is important to note that the results presented reflect the children's performance throughout one year of linguistic development and it is not possible to predict whether Italian will continue to be a strong or weak language. However, on the basis of the results obtained by comparing the two scales, it is possible to hypothesise that if the children who are developing Italian as a weak language continue to be exposed to a limited amount of input, their Italian will continue to be weak. This assumption has great implications for the study of bilingual development and it poses new questions that will need to be addressed in the future. If the input is limited and a language is constantly weak during childhood, can the child be considered a native speaker of that language? What implications does weakness have at different stages of linguistic development? Is the lack of input the main reason for a weak development or do other linguistic (and also non linguistic) factors come into play after the dominant language has become more established? Answering these questions is very important for the understanding of bilingual development not only in children, but also in adults. In a broader perspective, these questions concern the issue of language maintenance through the lifespan. If we assume that it is possible to go though different phases of linguistic weakness, what factors may affect the maintenance of the minority language through the years? Some studies have tested the proficiency of adult heritage language speakers and have found phenomena such as attrition, incomplete acquisition and language loss (Polinski 1997, Montrul 2004, Rothman 2007). Montrul (2004) uses the term "incomplete learners" to refer to adults who have failed to "completely" acquire the minority language. In L2 studies it is generally difficult to determine whether the adult has developed the minority language as a weak language since the earliest stages, or if there has been a progressive loss through the years, due to the contact with the majority language or to the lack of use of the minority one. The analysis of the adult backgrounds can only be assessed by administering questionnaires, and not by examining the parents' input. As Polinski (2008) observes, adult heritage speakers can represent the "crucial missing link" between L1 speakers, L2 learners and balanced bilinguals. It follows that children who develop a heritage language (like the children analysed in this thesis) represent another missing link between L1 and child L2 learners. Francesca, Paolo, Costanza and Matelda will go to English-speaking schools and, if the linguistic behaviour in their family is not

entirely modified, they are likely to be exposed to Italian on a daily basis. As the children grow up, more questions will emerge: are these children going to be identifiable as heritage speakers? What differences will we find between the children who developed Italian as a weak language and those who developed it as a strong language? What differences will we find between monolingual speakers and bilingual heritage speakers? These questions still need to be answered, and they can constitute the basis for further research.

7.3.1 Contribution to existing knowledge

Overall, this thesis presents three major findings, which can be summarised as follows:

- Italian can be considered a weak language if the child exhibits the following characteristics throughout the development: slow rate of acquisition, short MLU, presence of a high number of word order, agreement and verb inflection errors, limited lexicon, difficulty in making pragmatically correct subject selection (null vs. overt).
- Subject selection (overt – null) and subject position (preverbal – postverbal) are properties that can be tested to determine whether the child develops Italian as a weak or strong language.
- There is a relationship between quality and quantity of input and weak language development.

The main contributions provided by this study are represented by the methods used to test the weak language and the quality and quantity of input, which make it possible also to compare the results of the two different analyses on the basis of two scales. The employment of these methodologies allows us to determine whether Italian is a weak or strong language, and whether there is a relationship between input and weakness.

Another original contribution of this thesis lies in the exploration of interface properties of the language combining experimental and longitudinal data. The results of the analysis of postverbal subjects show that they are underused by bilingual children who develop Italian as the weak language. This result can be more widely interpreted as a difficulty to process structures requiring the activation of syntactic and pragmatic knowledge. This constitutes an important contribution to the growing

research on the acquisition of properties at interfaces. Both the experimental and the longitudinal results show that bilingual children who develop Italian as a weak language present difficulties in the production of postverbal subjects and direct object pronouns. This represents further evidence of the difficulty for "weak bilinguals" to process properties at the interface between syntax and pragmatics.

Overall, this study contributes to knowledge in different fields of the study of the bilingual language faculty, namely the weak language, the role of the input and the acquisition of interface properties. The hypotheses and the results presented in this thesis constitute a starting point for more in-depth research in these areas.

7.3.2 Limitations of this study and directions for future research

The areas explored in this thesis are quite diverse, and definitely require further careful and thorough analysis. Several issues need to be addressed. Firstly, this study provides an analysis of the weak language based on a new methodology, which should be further tested on larger populations. Moreover, the validity of the scales and values proposed to test the input and the weak language needs to be confirmed by employing larger amounts of data from monolingual and bilingual children.

There are important issues emerging from this thesis that are still unresolved and that should be addressed in future research. The first is the issue of attainment in 2L1 acquisition. In order to assess bilingual children, it is necessary to have more studies showing the differences between weak and strong language and showing how the two can be assessed. The analysis of the weak language could be expanded by including more factors and, by examining large amounts of data, it would be possible to determine which factors are more significant in assessing weakness. The same criterion applies to the analysis of the input. Future research should look more into the different characteristics of the input and possibly weigh the significance of each qualitative and quantitative aspect. Finally, more studies are needed to understand the relationship between child and adult heritage language, focusing on language maintenance, the role of the input and other factors which affect the ultimate attainment.

7.4 Conclusion

In this chapter I have summarised the main findings of my thesis and discussed the potential for future work. I have explained how this thesis constitutes a

methodological and theoretical contribution to knowledge in different areas of linguistics research and I have discussed some of the main limitations that have emerged and some of the areas that need to be further researched.

APPENDIX A

Li Wei’s classification of bilinguals

Li Wei's classification of bilinguals
(Li Wei 2000: 6-7)

Achieved bilingual	same as *late bilingual*.
Additive bilingual	someone whose two languages combine in a complementary and enriching fashion.
Ambilingual	same as *balanced bilingual*.
Ascendant bilingual	someone whose ability to function in a second language is developing due to increased use.
Ascribed bilingual	same as *early bilingual*.
Asymmetrical bilingual	see *receptive bilingual*.
Balanced bilingual	someone whose mastery of two languages is roughly equivalent.
Compound bilingual	someone whose two languages are learnt at the same time, often in the same context.
Consecutive bilingual	same as *successive bilingual*.
Co-ordinate bilingual	someone whose two languages are learnt in distinctively separate contexts.
Covert bilingual	someone who conceals his or her knowledge of a given language due to an attitudinal disposition.
Diagonal bilingual	someone who is bilingual in a non-standard language or a dialect and an unrelated standard language.
Dominant bilingual	someone with greater proficiency in one of his or her languages and uses it significantly more than the other language(s).
Dormant bilingual	someone who has emigrated to a foreign country for a considerable period of time and has little opportunity to keep the first language actively in use.
Early bilingual	someone who has acquired two languages early in childhood.
Equilingual	same as *balanced bilingual*.
Functional bilingual	someone who can operate in two languages with or without full fluency for the task in hand.
Horizontal bilingual	someone who is bilingual in two distinct languages which have a similar or equal status.
Incipient bilingual	someone at the early stages of bilingualism where one language is not fully developed.
Late bilingual	someone who has become a bilingual later than childhood.
Maximal bilingual	someone with near native control of two or more languages.
Minimal bilingual	someone with only a few words and phrases in a second language.
Natural bilingual	someone who has not undergone any specific training and who is often not in a position to

	translate or interpret with facility between two languages.
Passive bilingual	same as *receptive bilingual*.
Primary bilingual	same as *natural bilingual*.
Productive bilingual	someone who not only understands but also speaks and possibly writes in two or more languages.
Receptive bilingual	someone who understands a second language, in either its spoken or written form, or both, but does not necessarily speak or write it.
Recessive bilingual	someone who begins to feel some difficulty in either understanding or expressing him or herself with ease, due to lack of use.
Secondary bilingual	someone whose second language has been added to a first language via instruction.
Semibilingual	same as *receptive bilingual*.
Semilingual	someone with insufficient knowledge of either language.
Simultaneous bilingual	someone whose two languages are present from the onset of speech.
Subordinate bilingual	someone who exhibits interference in his or her language usage by reducing the patterns of the second language to those of the first.
Subtractive bilingual	someone whose second language is acquired at the expense of the aptitudes already acquired in the first language.
Successive bilingual	someone whose second language is added at some stage after the first has begun to develop.
Symmetrical bilingual	same as *balanced bilingual*.
Vertical bilingual	someone who is bilingual in a standard language and a distinct but related language or dialect.

APPENDIX B
Questionario Sul Contesto Linguistico Del Bambino Bilingue

Questionario Sul Contesto Linguistico Del Bambino Bilingue

DATI PERSONALI

1. Nome .....

2. Data di nascita .....

3. Luogo di residenza attuale .....

4. Precedenti luoghi di residenza
   a) Luogo ..... Dal ..... al .....
   b) Luogo ..... Dal ..... al .....
   c) Luogo ..... Dal ..... al .....

ALL'ASILO/CRÉCHE

5. Il bambino frequenta un asilo o crèche? .....

6. Se sì, da quando? .....

7. Quanti giorni alla settimana? .....

8. Quante ore alla settimana? .....

9. Ci sono stati cambiamenti di scuole-orari-giorni nel corso degli anni? Se sí, specificare.
.....
.....
.....

IN FAMIGLIA

10. Lingua madre della madre .....

11. Lingua madre del padre .....

12. Ore passate solo con la madre ogni settimana .....

13. Ore passate solo con il padre ogni settimana .....

14. Ore passate con entrambi .....
(se ci sono stati cambiamenti significativi durante gli anni, spiegare)
.....

.....

14b. Numero di ore di sonno al giorno .....
(se ci sono stati cambiamenti significativi durante gli anni, spiegare)

.....

.....

15. Che lingua parlano i genitori tra di loro? Se più d'una, spiegare in che contesti e con quale frequenza vengono usate le diverse lingue.
.....

.....

.....

16. Che lingua parla il padre con il bambino?
.....

17. Che lingua parla la madre con il bambino?
.....

18. Quale lingua parla prevalentemente il bambino?.....

19. Per quante ore in media al giorno il bambino è esposto all'italiano? (includere il tempo passato guardando la televisione in italiano o ascoltando persone che parlano italiano) .....

20. Per quante ore in media al giorno il bambino è esposto all'inglese? (e/o ad altre lingue?) .....

21. In quali dei seguenti casi la madre si rivolge al figlio nella propria lingua madre? (è possibile scegliere più di una risposta)

- ☐ sempre
- ☐ quando sono soli
- ☐ in casa
- ☐ nel paese in cui la lingua è parlata
- ☐ in presenza di persone che parlano la lingua
- ☐ raramente
- ☐ mai

☐ altro (specificare)

.....

.....

22. In quali dei seguenti casi il padre si rivolge al figlio nella propria lingua madre? (è possibile scegliere più di una risposta)

☐ sempre
☐ quando sono soli
☐ in casa
☐ nel paese in cui la lingua è parlata
☐ in presenza di persone che parlano la lingua
☐ raramente
☐ mai
☐ altro (specificare)

.....
.....

23. Con quale frequenza il padre parla la lingua della madre?

☐ sempre
☐ spesso
☐ qualche volta
☐ raramente
☐ mai

24. In quali contesti?

.....
.....

25. Con quale frequenza la madre parla la lingua del padre?

☐ sempre
☐ spesso
☐ qualche volta
☐ raramente
☐ mai

26. In quali contesti?

.....
.....

27. Con la madre, il bambino parla:

☐ sempre la lingua della madre
☐ quasi sempre la lingua della madre
☐ a volte una lingua, a volte l'altra
☐ quasi mai la lingua della madre

☐ mai la lingua della madre

28. Con il padre, il bambino parla:
- ☐ sempre la lingua del padre
- ☐ quasi sempre la lingua del padre
- ☐ a volte una lingua, a volte l'altra
- ☐ quasi mai la lingua del padre
- ☐ mai la lingua del padre

29. Oltre ai genitori, con chi altro il bambino parla italiano?
.....

.....

30. Con quale frequenza?
.....

.....

31. Quanto tempo il bambino ha passato in Italia da quando è nato?
1° anno.....
2° anno.....
3° anno.....
4° anno.....

32. Se le visite sono frequenti, in media, ogni anno, il bambino quanto tempo passa in Italia? .....

DOMANDE PER IL GENITORE ITALIANO

33. Il tuo bambino...
- ☐ non capisce l'italiano
- ☐ non so se capisce l'italiano
- ☐ a volte capisce, a volte no
- ☐ capisce tutto, ma non parla
- ☐ mi capisce quando parlo italiano, ma risponde sempre in inglese
- ☐ mi capisce quando parlo italiano e a volte risponde in italiano
- ☐ mi capisce quando parlo italiano e mi risponde sempre in italiano

34. Vuoi che il tuo bambino impari ad usare l'italiano?
- ☐ sì, per me è importantissimo che mio figlio impari l'italiano
- ☐ sì, per me è importante che mio figlio impari l'italiano
- ☐ spero che lo impari perchè potrebbe essere utile
- ☐ voglio che lo impari, ma l'inglese è una lingua più utile
- ☐ l'importante è che impari l'inglese, poi si vedrà
- ☐ non importa se non impara l'italiano

35. Il tuo coniuge vuole che il bambino impari ad usare l'italiano?

- ☐ sì, per lui/lei è importantissimo che nostro figlio impari l'italiano
- ☐ sì, per lui/lei è importante che nostro figlio impari l'italiano
- ☐ spera che lo impari perchè potrebbe essere utile
- ☐ vuole che lo impari, ma l'inglese è una lingua più utile
- ☐ l'importante è che impari l'inglese, poi si vedrà
- ☐ non importa se non impara l'italiano

36. In quali modi in famiglia si cerca di promuovere l'uso della lingua italiana?

.....

.....

.....

.....

.....

.....

Annotazioni

Questa sezione è dedicata ad annotazioni o commenti che i genitori ritengono rilevanti per capire meglio il contesto linguistico in cui vive il bambino, dettagli che possono risultare utili per lo studio del suo sviluppo linguistico o informazioni che non sono emerse dal questionario.

.....

.....

.....

.....

.....

.....

APPENDIX C
Questionnaire on the Linguistic Background of the Bilingual Child

Questionnaire on the Linguistic Background of the Bilingual Child

PERSONAL INFORMATION

1. Name .....

2. Date of birth .....

3. Place of current residence .....

4. Previous places of residence
   a) Place ..... From ..... to .....
   b) Place ..... From ..... to .....
   c) Place ..... From ..... to .....

DAYCARE/CRÉCHE

5. Does your child attend a daycare centre/crèche? .....

6. If so, since when? .....

7. How many days a week? .....

8. How many hours a week? .....

9. Have there been changes in school-times-days during the years? If so, specify.
.....

.....

.....

IN THE FAMILY

10. Mother's language .....

11. Father's language .....

12. Hours spent only with the mother each week .....

13. Hours spent only with the father each week .....

14. Hours spent with both .....
(if there have been significant changes, specify)
.....

.....

14b. Number of hours of sleep a day .....
(if there have been significant changes, specify)
.....

.....

15. What language do the parents speak to each other? If it is more than one, explain in which contexts and how frequently the two languages are used.
.....

.....

.....

16. What language does the father speak with the child?
.....

17. What language does the mother speak with the child?
.....

18. Which language does the child speak predominantly ?
.....

19. How many hours a day on average is the child exposed to Italian? (include the time spent watching Italian TV or listening to people speaking Italian)
.....

20. How many hours a day on average is the child exposed to English? (and/or other languages?)
.....

21. In which of the following cases does the mother address the child in her native language? (you can choose more than one answer)

- ☐ always
- ☐ when they are alone
- ☐ at home
- ☐ in the country where the language is spoken
- ☐ in the presence of people who speak the language
- ☐ rarely
- ☐ never
- ☐ other (specify)

  .....

  .....

22. In which of the following cases does the father address the child in his native language? (you can choose more than one answer)

- ☐ always
- ☐ when they are alone
- ☐ at home
- ☐ in the country where the language is spoken
- ☐ in the presence of people who speak the language
- ☐ rarely
- ☐ never
- ☐ other (specify)

  .....

  .....

23. How often does the father use the mother's language?

- ☐ always
- ☐ often
- ☐ sometimes
- ☐ rarely
- ☐ never

24. In which contexts?

.....

.....

25. How often does the mother use the father's language?

- ☐ always
- ☐ often
- ☐ sometimes
- ☐ rarely
- ☐ never

26. In which contexts?

.....

.....

27. With the mother, the child speaks:

- ☐ always the mother's language
- ☐ almost always the mother's language
- ☐ sometimes one language, sometimes the other
- ☐ hardly ever the mother's language
- ☐ never the mother's language

28. With the father, the child speaks:

- ☐ always the father's language
- ☐ almost always the father's language
- ☐ sometimes one language, sometimes the other
- ☐ hardly ever the father's language
- ☐ never the father's language

29. Excluding the parents, who else does the child speak Italian to?

.....

.....

30. How often?

.....

.....

31. How much time did the child spend in Italy since birth?

1st year.....
2nd year.....
3rd year.....
4th year.....

32. If the trips are frequent, on average, how much time does the child spend in Italy each year?

.....

QUESTIONS FOR THE ITALIAN PARENT

33. Your child...

- ☐ does not understand Italian
- ☐ I don't know if he/she understands Italian
- ☐ Sometimes he/she understands, sometimes he/she does not
- ☐ He/she understands everything, but does not talk
- ☐ He/she understands me when I speak Italian, but answers in English
- ☐ He/she understands me when I speak Italian, and sometimes answers in Italian
- ☐ He/she understands me when I speak Italian, and always answers in Italian

34. Do you want your child to learn to use Italian?

- ☐ Yes, for me it is very important that my child learn Italian
- ☐ Yes, for me it is important that my child learn Italian
- ☐ I hope he/she learns because it could be useful
- ☐ I want him/her to learn Italian, but English is more useful
- ☐ It is important that he/she learns English, then we'll see
- ☐ It is not important that he/she learns Italian

35. Does your partner want your child to use Italian?

- ☐ Yes, for him/her it is very important that my child learn Italian
- ☐ Yes, for him/her it is important that my child learn Italian
- ☐ He/She hopes that the child learns because it could be useful
- ☐ He/She wants him/her to learn Italian, but English is more useful
- ☐ It is important that he/she learns English, then we'll see
- ☐ It is not important that he/she learns Italian

36. In what ways is your family trying to promote the use of Italian?

.....

.....

.....

.....

.....

.....

Annotations

This section is dedicated to annotations or comments that the parents consider relevant to have a better understanding of the child's linguistic contexts, of details that can be useful for the study of language development or to provide information that has not emerged through the questionnaire.

.....

.....

.....

.....

.....

.....

APPENDIX D

Values used to measure the input

Values used to measure the input

Table D.1 Quantity of input

Values	Quantity
1	10-14
1.5	15-19
2	20-24
2.5	25-29
3	30-34
3.5	35-39
4	40-44
4.5	45-49
5	50-54
5.5	55-59
6	60-64
6.5	65-69
7	70-74
7.5	75-79
8	80-84
8.5	85-89
9	90-94
9.5	95-99
10	100

Table D.2 Output

Values	Output
1	100-149
1.5	150-199
2	200-249
2.5	250-299
3	300-349
3.5	350-399
4	400-449
4.5	450-499
5	500-549
5.5	550-599
6	600-649
6.5	650-699
7	700-749
7.5	750-799
8	800-849
8.5	850-899
9	900-949
9.5	950-999
10	1000

Table D.3 Vocabulary

Values	Vocabulary
1	25-49
1.5	50-74
2	75-99
2.5	100-124
3	125-149
3.5	150-174
4	175-199
4.5	200-224
5	225-249
5.5	250-274
6	275-299
6.5	300-325
7	325-349
7.5	350-374
8	375-399
8.5	400-425
9	425-449
9.5	450-474
10	475-479

Table D.4 MLU

Values	MLU
1	4.0-4.4
1.5	4.5-4.9
2	5.0-5.4
2.5	5.5-5.9
3	6.0-6.4
3.5	6.5-6.9
4	7.0-7.4
4.5	7.5-7.9
5	8.0-8.4
5.5	8.5-8.9
6	9.0-9.4
6.5	9.5-9.9
7	10.0-10.4
7.5	10.5-10.9
8	11.0-11.4
8.5	11.5-11.9
9	12.0-12.4
9.5	12.5-12.9
10	13.0-13.4

APPENDIX E

Number of Target Deviant Forms

Number of Target Deviant Forms

Table E.1 Target deviant forms - Matelda

Age	Word Order	Gender Agreement	Number Agreement	Verb morphology
2;6.23	0	0	0	0
2;8.11	0	0	1	1
2;9.07	0	0	1	4
2;10.12	0	0	0	1
3;1.15	0	1	0	0
3;2.20	0	0	0	0

Table E.2 Target deviant forms - Paolo

Age	Word Order	Gender Agreement	Number Agreement	Verb morphology
3;1.27	0	0	0	1
3;3.23	2	2	0	0
3;4.25	0	1	0	0
3;7.10	1	0	0	0
3;10.19	2	1	0	6
3;11.17	1	2	0	1
4;0.29	1	1	0	8
4;1.28	2	5	0	3

Table E.3 Target deviant forms - Francesca

Age	Word Order	Gender Agreement	Number Agreement	Verb morphology
2;4.20	0	0	0	0
2;5.10	0	1	0	3
2;6.19	0	3	0	2
2;7.28	1	1	1	0
2;9.07	0	1	0	1
2;10.17	1	3	0	1
3.0.17	1	1	1	3
3;1.17	0	0	0	10
3;2.27	2	2	0	6
3;5.0	0	1	1	9

Table E.4 Target deviant forms - Costanza

Age	Word Order	Gender Agreement	Number Agreement	Verb morphology
1;1.16	0	0	0	0
1;12.10	0	1	0	2
2;2.17	0	0	0	2
2;4.09	0	0	0	0
2;6.07	1	0	0	2
2;7.16	0	0	0	0
2;9.14	0	0	0	0

APPENDIX F

Lexical production in Italian, classified by age and word category

Lexical production in Italian, classified by age and word category

Table F.1 Lexical production - Matelda

Age	Nouns	Adjectives	Verbs
2;6.23	cane papi passeggiata puzzle		va
2;8.11	aereo bambina capriole carota copertina dentini faccia fogliolina patate prosciutto scarpe spinaci zio	marrone nuova viola	andiamo colori è fatto finito funziona portato senti
2;9.7	aereo amico animali biscotti borsa braccia cammello casa ciabatta collo erba fine fiocco gambe ghiaccio giraffa libretto mela miele negozio neve nonna occhiali orsetto palloncini personaggio	corti grande neri nuovo piano piccolina ricciolini rosa rosso solo tante verde	caduto, cade devi dice faccio vedere fai, facciamo, ho fatto finito guarda ho messo, messa ho preso, prende mangia pennella sembra sono sta diventando tolgo va, sono andate viene

	pompetta pongo serpente signore stella storie, storia tigre torta		
2;10.12	amichetto bagno bocca capelli dentini lavoretto letto mosca nasetto ponte scuola settembre	bianco blu nero piccino sola	apri dice dormire faccio, fai, fare, ha fatta guarda puoi scende togli usare vado vedo viene vuole, vuoi, voglio
3;1.15	ape bambina barca becco fiore gola miele musica pulcino	alto blu meglio piccolo	aspetta bagnamo fare, facciamo, sta facendo finito guarda hai visto lasciamo mettiamo pizzica puoi sai si è appiccicato si è attaccato si muove voglio
3;2.20	coda, codina corna elefante, elefanti gamba gonna mucca piscina trottola	altro bello bianco blu	bagnato, bagno, bagnata è caduto facciamo guarda ho fatto mangiata, mangia, mangio mi serve

			nuotare prendiamo proviamo si attaccano so

Table F.2 Lexical production - Paolo

Age	Nouns	Adjectives	Verbs
3;1.27	bambino colore libro	cattivo caldo giallo	fa vieni
3;3.23	cammello cane casa fuoco gambe leone lupo mano onde scavatore mare balena bocca acqua occhio cuore	bello brutto buono cattivo forte nuovo piccolo rosso tutto	è morto fatto guarda morto scappa sono
3;4.25	neve latte buio luna lupo casa amici maialino gamba nonna	lontano piccolo	guarda mangiare fa
3;7.10	cavalla mani sole notte sera soldi burrone		rotto scotta chiudo è
3;10.19	acqua anatre, anatroccoli cacciatore casa coniglio farfalla gallina mucche	brutto buono cattivo grande grossa	camminare da diventa è è andato finito ha mangiare

	pappagallo pugno tigre uovo		prendi, prendono ride sai scritto sembra spara viene voglio
3;11.17	asilo braccio calzino cane capelli colore drago elefante, elefanti faccia fronte gambe lavoro occhi, occhio occhiali pompiere rumore tappi	altro, altra brutto giallo grandi, grande nero piccolo rossi, rosso verde, verdi	aiuta butta colorare è, sono fare, fa hanno leggi posso so, sai spegne vai voglio
4;0.29	braccio cacciatore cane coda coniglio denti dottore frutta leone mano matite pancia pane rocce tigre uomo volpe	bello cattivo grande italiano lontano lungo tutta	aiuta, aiuti corri, corre esce fai, fa, farlo guarda legge, leggi mangia, mangiarlo messa passi piangi rotto saltare serve so sono, è, è stato sparato uccide, uccido vado vieni voglio, vuole
4;1.28	albero carote	brutto nostra	dormi è

	casa colori computer forbice leone lotta occhio, occhi serpente	tanti	fare, fa, fai, fanno, faccio guarda mettere prendi si rompe, si è rotto sta voglio

Table F.3 Lexical production - Francesca

Age	Nouns	Adjectives	Verbs
2;4.20	anatre bagno cavallo		vuoi
2;5.10	bambina giraffe palla	bianco rosa verde	
2;6.19	coccodrillo forchetta fuoco lavoro luce occhi porcellino	altro sporchi tutti	apri buttare fai finito lascia siediti vuoi
2;7.28	acqua albero calze coccodrillo colore foglie gatto macchina mela pera pesce rana, rane sole	giallo mio nero, nera	si accende scappato vuoi
2;9.07	acqua asino autobus bocca buio casa foca giacca luce lupo macchina motorino nonna palla passeggino pesce porta torta	bagnata mia	apre, apri dorme è fa, fai guarda leggi metti sta va vuoi

2;10.17	acqua aereo bambina, bambini, bambino banana casa conchiglia cucina formaggio frutta macchine motorino pane pesce pizza pollo	mio piccola, piccolo rosso tua	andare colorare è fai hai letto leggi mangio messa qui va vuoi
3;0.17	acqua cagnolino festa formaggio gattino luna mucca pecora porcellino sole spesa uovo	altro	andare è, sono fare ha portato leggi messi so tieni vedi
3;1.17	acqua amici baci cagnolino, cane casa castello festa libro mare nonno occhi orecchie orsetto, orso pesce piedi porcellino sole storia telefono	altro, altra chiuso grande inglese mio piccolo tuo tutto	apre, aprilo, apro attacca è fai, fare, fatto, fai finita guarda ha leggi, legge mangi metti mi piace prendi regala sta trovare va vuoi

	testa torta vestito		
3;2.27	albicocca amica autobus bambini bicchiere bocche braccio caffè calzini cane capelli cappello colori figlia fiori foglie gatto lavoro letto limone luce mandarino mare mucca occhi ombrello palloncini parco-giochi pesca pezzo pile porcellini scale scarpe scimmiette sole tavolo telecomando uccello zoo	altro arancione attenta blu facile giallo grande mia, mie nero piccolo rosa rosso tua uguale viola	andiamo cadere colorare, colora devi dormire è, sono fai, faccio, fa, fatto finito funziona giocare, giochiamo gira guarda guidi, guida, guido hai, ho lavorare leggi mangiare metti, mette mi piace, mi piacciono, ti piace, ti piacciono prende, prendi rilassati salta scotta se l’è messi siediti spegne, spegni sta, stare tieni toccare togliermi torno usi va, vado, vai vieni vuoi
3;5.0	borsa burro cereali chiavi cosa	altri bello, bella freddo grande mia, mio	andata, andiamo apri, apro baci è faccio, facciamo, fa

	fiore giardino macchina madre mani motorino palla parti passeggiata porta posto principessa schifo sedia spesa uccellino	sporco	giocano, giocare, giochiamo guarda guidi ho fatto lascia, lasciamo mangiamo piove prendi provi scaviamo si arrabbia sono va, vai vedere vuoi, voglio

Table F.4 Lexical production - Costanza

Age	Nouns	Adjectives	Verbs
1;1.16	leone giraffa lavoro moto piedi jeans calzine bambolina cappello	calda	apri hai messo accende, accendi
1;12.10	aereo bimba canzoncina nonna uccellino	bella femmina	è tieni
2;2.17	acqua albero asinello capriole cervi febbre gallo lupo maiale scarpe scheletro sete signora zio, zia	bianco, bianchi gialla piccolo, piccola rossi triste	aspetta balla dammi, dai è fa, fare, fai ha, ha avuto, ho ho trovato leggi messa nascosto piange scrivo voglio
2;4.09	acqua bambolina becchino bimba cacca cavallo copertina cuore fame fazzoletto fermaglio forma formaggio frutta fuoco gallinella macchina	altro bagnata bella buona duro piccolo suo	aprire asciugare attenta bada brucia cambiare casca comprare correre devo, devi, deve dorme è fa, fai, fare guarda hai finito, ho finito ho lascia

	maglietta nanna ovetto pappa pasta patate pietra pipí salsiccia scalino scarpe soldino vasino		mangia, mangiare metti, si mette premere prendere scotta vado vieni vuoi, voglio
2;6.07	banana biscotto blocchetto casa colore fragola gatto gioco giungla lupi piede scarpe telefonino triciclo yogurt	azzurre comode giallo preferita rosa rotto tuo	ballo devo, deve è facciamo, sto facendo mangia mettere saltiamo sta tieni va vedere viene voglio
2;7.16	acqua antenne becco bocca capriole castello coniglio farfalla, farfallona gatto ippopotamo legno lupo manina mare mucca porta regalo storia	piccolo rosso sua	apre dai dice è fa giocare grugnisce guarda leggere, leggi mangia nitrisce scritto staccata starnazza taglia vedere vuoi, voglio

2;9.14	bacino bamboline, bambolina bimba, bimbo cacca caldo (N) casa cioccolata coccinella cuoricino farfalla fiore fratellino freddo (N) giardino ginocchio giornalino mandorle mattonella negozio palla piedi pipi porta sandalini scarpe sorellina tenda uccellino volta	altri buono grande inglese italiano mio piccola rosa solo sua tutti ultima viola	andare, andiamo aspetti cade, cadi camminare, cammina compriamo cucino dorme e’ faccio, facciamo fare giochiamo, gioco ho trovato lavorare leggo mantengo metti, metto mi alzo mi chiamo mi siedo, siediti mi sto nascondendo, nascondiamo, nasconditi parla, parlano piange possiamo, possono puzza sai si apre si sveglia sono sto mangiando trova usciamo uscire vado vedere, vedi, vedo vieni vogliamo

APPENDIX G
Values for the assessment of the weak language

Values for the assessment of the weak language

Table G.1 Scale

Age	Scale	Article Om. %	Scale	Vocab	Scale	Target Dev. %	Scale	Subjects	Scale
24-25-26	10	0-9	10	100	10	0-0.9	10	96-100	10
27-28-29	9	10-19	9	90-99	9	1-1.9	9	91-95	9
30-31-32	8	20-29	8	80-89	8	2-2.9	8	86-90	8
33-34-35	7	30-39	7	70-79	7	3-3.9	7	81-85	7
36-37-38	6	40-49	6	60-69	6	4-4.9	6	76-80	6
39-40-41	5	50-59	5	50-59	5	5-5.9	5	71-75	5
42-43-44	4	60-69	4	40-49	4	6-6.9	4	66-70	4
45-46-47	3	70-79	3	30-39	3	7-7.9	3	61-65	3
48-49-50	2	80-89	2	20-29	2	8-8.9	2	56-60	2
51-52-53	1	90-99	1	10-19	1	9-9.9	1	51-55	1
54-55-56	0	100	0	0-9	0	10	0	64-50	0

APPENDIX H
Animali in casa

Animali in casa

Text

Sul tavolo ci sono una mela, dei cereali, delle carote e un succo di frutta.
On the table there are an apple, some cereal, some carrots and an orange juice.

Arriva un ragno e mangia la mela.
A spider arrives and eats the apple.

La mamma arriva e dice: Oh!(Child's name) Che fine ha fatto la mela?
The mum arrives and says: Oh! What happened to the apple?/Where's the apple gone?

TARGET ANSWERS
L'ha mangiata il ragno.
L'ha mangiata un ragno.

Arriva una gallina e mangia i cerali.
A hen arrives and eats the cereal.

La mamma arriva e dice: Oh! (Child's name) Che fine hanno fatto i cereali?
The mum arrives and says: Oh! What happened to the ceral?

TARGET ANSWERS
Li ha mangiati la gallina.
Li ha mangiati una gallina.

Arriva un coniglio e mangia le carote.
A rabbit arrives and eats the carrots.

La mamma arriva e dice: Oh! (Child's name) Che fine hanno fatto le carote?
The mum arrives and says: What happened to the carrots?

TARGET ANSWERS
Le ha mangiate la gallina.
Le ha mangiate una gallina.

Arriva una rana e beve il succo di frutta.
A frog arrives and drinks the orange juice.

La mamma arriva e dice: Oh! (Child's name) Che fine ha fatto il succo di frutta?
The mum arrives and says: Oh! What happened to the orange juice?

TARGET ANSWERS
Lo ha bevuto la rana/L'ha bevuto la rana.
Lo ha bevuto una rana/L'ha bevuto una rana.

H.1 Stills from the animation

Figure 2.1

Figure 2.2

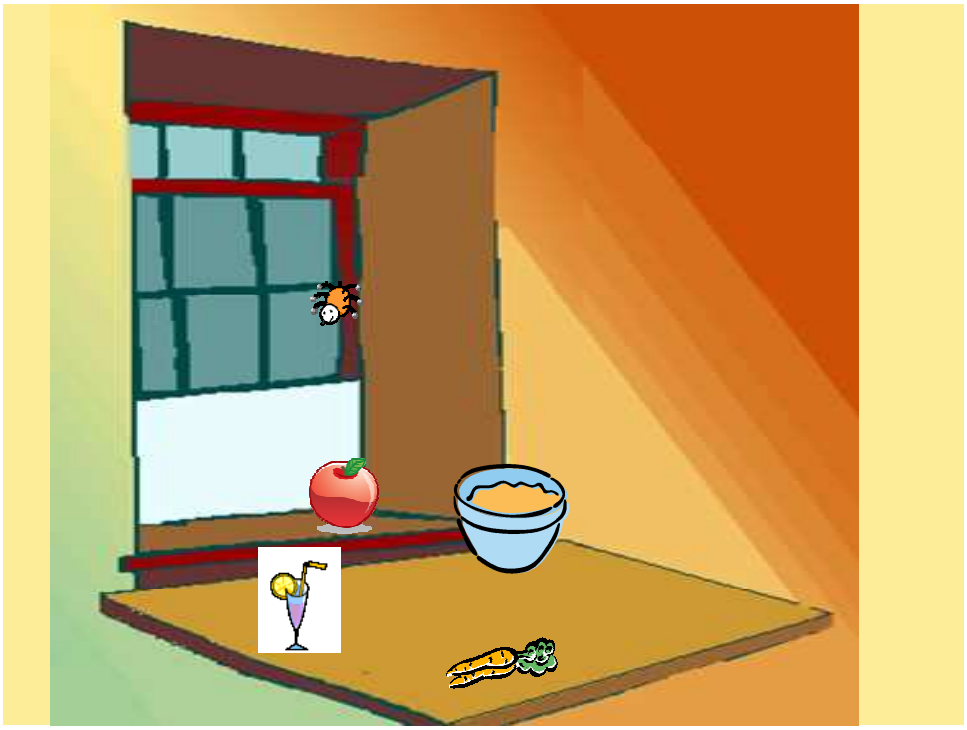


Figure 2.3

Figure 2.4

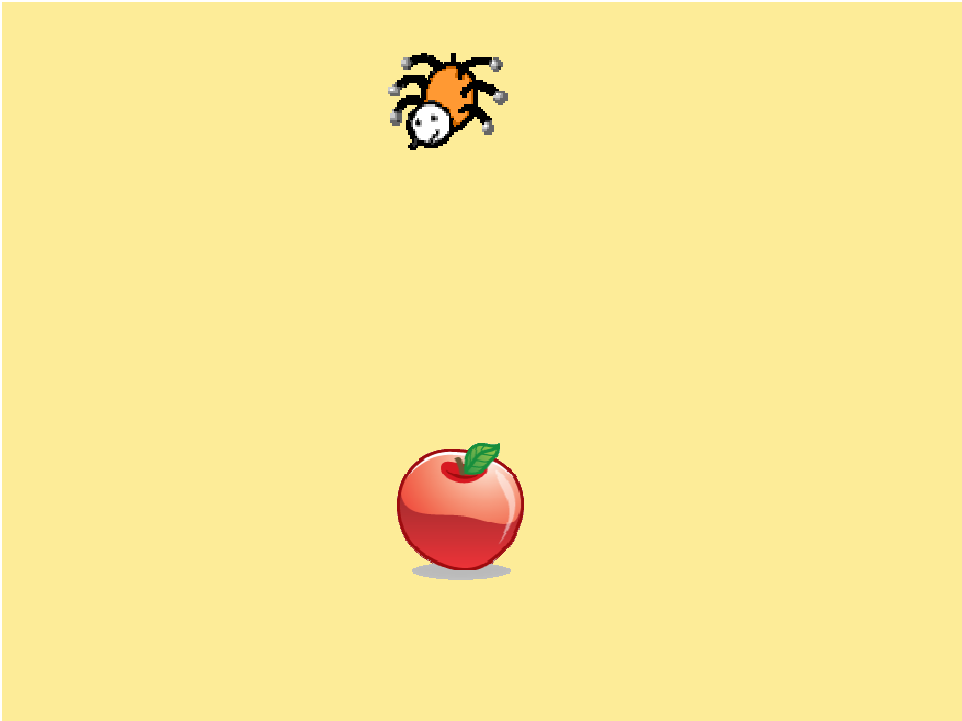


Figure 2.5

Figure 2.6

Figure 2.7

Figure 2.8

Figure 2.9

Figure 2.10

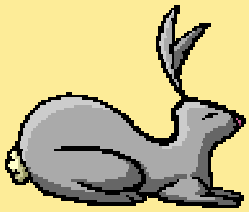


Figure 2.11

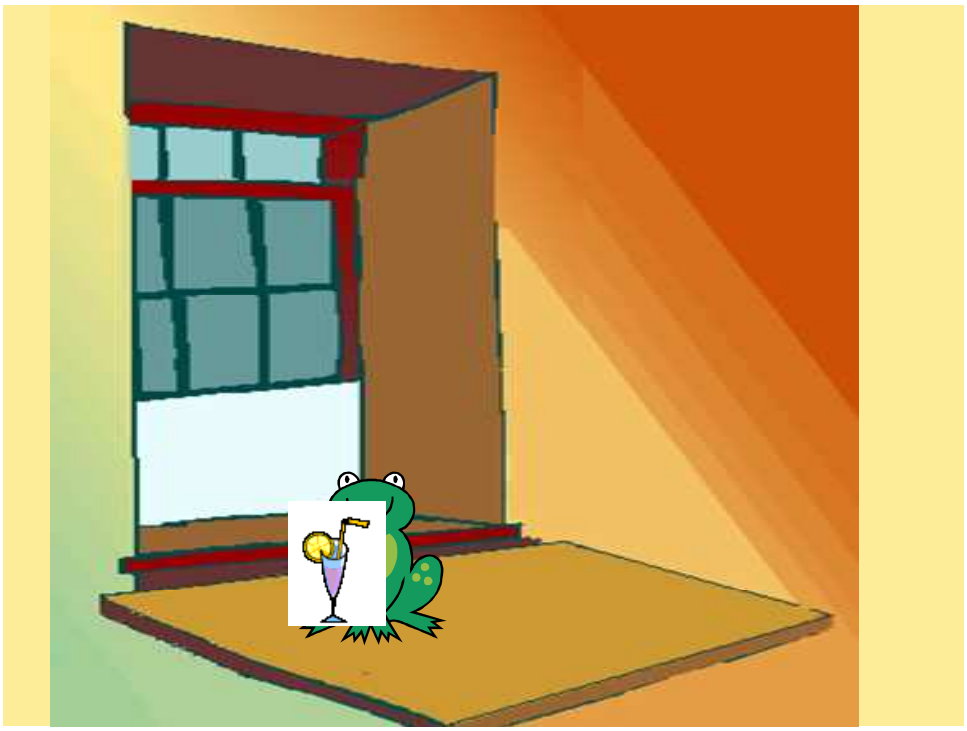


Figure 2.12

Figure 2.13

References

Adragão, M. and Costa, J. 2003. On the status of preverbal subjects in null subject languages: evidence from acquisition. *Proceedings of GALA* 2003, Vol. 1.

Akhtar, N., and Tomasello, M. 2000. The social nature of words and word learning. In Golinkoff et al. (eds.). *Becoming a word learner: A debate on lexical acquisition*. Oxford: Oxford University Press. pp115–135.

Almgren, M. and Barreña, A. 2001. Bilingual acquisition and separation of linguistic codes: ergativity in Basque versus accusativity in Spanish. In Nelson, K.E., Aksu-Koç, A., Johnson, C. *Children's Language*. Volume 11, pp26-48.

Antelmi, D. 1997. *La prima grammatical dell'italiano: indagine longitudinale sull'acquisizione della morfosintassi italiana*. Bologna: il Mulino.

Argyri, E. and Sorace, A. 2007. Crosslinguistic influence and language dominance in older bilingual children. *Bilingualism: Language and Cognition*. 10 (1), pp79-99.

Arias, J., Kintana, N., Rakow, M. and Rieckborn, S. 2005. Sprachdominanz: Konzepte und Kriterien. *Working Papers on Multilingualism*. University of Hamburg.

Avram, L. and Coene, M. 2003. What early clitics can tell us about early subjects. *Proceedings of GALA, Generative Approaches to Language Acquisition*. pp93-102.

Avrutin, S. 1999. *Development of the Syntax-Discourse Interface*. Kluwer, Dordrecht.

Baldwin, D. 2000. Interpersonal understanding fuels knowledge acquisition. *Current Directions in Psychological Science*. 9, pp40-45.

Behrens, H. 2006. The input-output relationship in first language acquisition. *Language and Cognitive Processes*. 21 (1-3), pp2-24.

Belletti, A. 1988. The Case of Unaccusatives. *Linguistic Inquiry*. 19 (1), pp1-34.

Belletti, A. 2001. "Inversion" as Focalization. In Hulk, A., Pollock, J.Y. (eds.) *Subject Inversion in Romance and the Theory of Universal Grammar*, Oxford University Press. pp60-89.

Belletti, A. and Leonini, C. 2004. Subject inversion in L2 Italian. Foster Cohen, S. Sharwood, M. Sorace, A. Ota, M. (eds.), *Eurosla Yearbook*, J. Benjamin, pp95-118.

Belletti, A., Bennati, E. and Sorace, A. 2007. Theoretical and developmental issues in the syntax of subjects: evidence from near-native Italian. *Natural Language and Linguistic Theory*. 25, pp657-689.

Bernardini, P. 2004. *L'Italiano come prima e seconda (madre)lingua. Indagine longitudinale sullo sviluppo del DP*. Lunds Universitet.

Bernardini, P. and Schlyter, S. 2004. Growing syntactic structure and code-mixing in the weaker language: the Ivy Hypothesis. *Bilingualism: Language and Cognition*. 7 (1), pp49-69.

Bettoni, C., Di Biase, B. and Nuzzo, E. (2009), Postverbal subject in Italian L2 – a Processability Theory approach. In Keßler, J.-U. and Keatinge, D. (eds.), *Research in Second Language Acquisition: Empirical Evidence Across Languages*, Newcastle upon Tyne, Cambridge Scholars.

Birdsong, D. 1999. Whys and why nots of the Critical Period Hypothesis. In *Second language acquisition and the Critical Period Hypothesis.* Birdsong, D. (ed.). Mahwah, NJ: Erlbaum. pp1–22.

Bonnesen, M. 2009. The status of the "weaker" language in unbalanced French/German bilingual language acquisition. *Bilingualism: Language and Cognition*. 12 (2), pp177-192.

Bornstein , M.H., Haynes, M.O, and Painter, K. M. 1998. Sources of child vocabulary competence: a multivariate model. *Journal of Child Language*. 25, pp367-393.

Borovsky, A., and Elman, J.L. 2006. Language input and semantic categories: A relation between cognition and early word learning. *Journal of Child Language*. 33, pp759-790.

Bosch, L. and Sebastián-Gallés, N. 1997. Native-language recognition abilities in four-month-old infants form monolingual and bilingual environments. *Cognition*. 65, pp33-69.

Bosch, L., Sebastián-Gallés, N. 2001. Early language differentiation in bilingual infants. In Cenoz, J. and Genesee, F. (eds.), *Trends in bilingual acquisition*. John Benjamins. pp71-93.

Bottari, P., Cipriani, P., Chilosi, A. M. Pfanner, L. 2001. The Italian determiner system in normal acquisition, specific language impairment, and childhood aphasia. *Brain and Language*. 77 (3), pp283-293.

Brown, R. 1973. *A first language: the early stages*. Cambridge, Mass.: Harvard University Press.

Burkhardt, P. 2005. *The syntax-discourse interface. Representing and interpreting dependency*, John Benjamins.

Cabré Sans, Y. and Gavarró, A. 2007. Subject distribution and verb classes in child Catalan. *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA).* Cascadilla Proceedings Project, Somerville, MA, USA

Cantone, K. 2007. *Codeswitching in bilingual children*. Springer.

Cantone, K. 2008. Un nase or una nase? What gender marking within switched DPs reveals about the architecture of the bilingual language faculty. *Lingua*. 118 (6), pp810-826.

Cantone, K., Kupisch, T., Müller, N. and Schmitz, K. 2008. Rethinking language dominance in bilingual children. *Linguistische Berichte*. 215, pp300-343.

Caselli, C., Casadio, P. and Bates, E. 1999. A comparison of the transition from first words to grammar in English and Italian. *Journal of Child language*. 26, pp69-111.

Caselli, M. C. and Casadio, P. 1995. *Il primo vocabolario del bambino. Guida all'uso del questionario MacArthur per la valutazione della comunicazione e del linguaggio nei primi anni di vita.* Franco Angeli, Milano.

Caselli, M. C., Pasqualetti, P. and Stefanini, S. 2007. *Parole e frasi nel "Primo vocabolario del bambino". Nuovi dati normativi fra 18 e 36 mesi e Forma breve del questionario.* Franco Angeli, Milano.

Cecchetto, C. and Rizzi, L. 2000. A naturalistic approach to language. In Nannini, S. and Sandkühler, H.J. (eds.) *Naturalism in the cognitive sciences and in the philosophy of mind*. Peter Lang Europäischer Verlag der Wissenschaften. pp117-130.

Central Statistics Office. 2006. *Non Irish Nationals living in Ireland*. Stationery Office, Dublin, Ireland.

Chomsky, N. 1975. *Reflections on language*. Pantheon.

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

Cipriani, P., Chilosi, A., Bottari, P. and Pfanner, L. 1993. *L'acquisizione della morfosintassi: fasi e processi*, Unipress, Padova.

Clahsen, H., Penke, M. and Parodi, T. 1993/1994. Functional categories in early child German. *Laguage acquisition*. 3, pp395-429.

Cook, V. and Newson, M. 2007. *Chomsky's Universal Grammar*. Wiley-Blackwell.

Costa, J. 2001. Postverbal subjects and agreement in unaccusative contexts in European Portuguese. *The Linguistic Review*. 18, pp1–17.

Curtiss, S. 1977. *Genie: a psycholinguistic study of a modern-day "wild child"*. Boston: Academic Press.

D'Odorico, L. 2001. Vocabulary development in Italian children: a longitudinal evaluation of quantitative and qualitative aspects. *Journal of Child Language.* 28 (2), pp351-372.

De Houwer, A. 1990. *The acquisition of two languages from birth: a case study.* Cambridge: Cambridge University Press.

De Houwer, A. 1995. Bilingual language acquisition. In Fletcher, P. and MacWhinney, B. (eds.) *The handbook of child language*. Cambridge: Blackwell. pp219-250.

De Houwer, A. 2007. Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*. 28 (3), pp411-424.

De Houwer, A. 2009. *Bilingual first language acquisition*. Multilingual Matters.

Deuchar, M. and Quay, S. 1998. One vs. two systems in early bilingual syntax: two versions of the question. *Bilingualism: Language and Cognition*. 1 (3), pp231-243.

Devescovi, A. and Caselli, M. C. 2007. Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language & Communication Disorders*. 42 (2), pp187-208.

Döpke, S. 1992. Approaches to first language acquisition: evidence from simultaneous bilingualism. *Australian Review in Applied Linguistics*. 15 (2), pp137–150.

Döpke, S. 1998. Competing language structures: the acquisition of verb placement in bilingual German-English children. *Journal of Child Language*. 25 (3), pp555-584.

Döpke, S. 2000. *The weaker language in simultaneous bilingualism: why it is not like L2*. Unpublished Manuscript, Monash University.

Ferrari, I. and Matteini, S. 2008. On the residual omission of determiners in Italian L1: a case study. *STiL – Studies in Linguistics Vol.2 - CISCL Working Papers*, University of Siena.

Genesee, F. 1989. Early bilingual development: one language or two? *Journal of Child Language*. 16 (1), pp161-179.

Genesee, F., Nicoladis, E., and Paradis, J. 1995. Language differentiation in early bilingual development. *Journal of Child Language*. 22, pp611-631.

Gleitman, L.R., Newport, E.L. and Gleitman, H. 1984. The current status of the motherese hypothesis. *Journal of Child Language*. 11 (1), pp43-79.

Grinstead, J. 2004. Subjects and interface delay in child Spanish and Catalan. *Language*. 80 (1), pp40-72.

Grosjean F. 1982. *Life with two languages: An introduction to bilingualism*. Harvard University. Press, Cambridge, Mass.

Grosjean, F. 2008. *Studying bilinguals*. Oxford: Oxford University Press.

Guasti M. T. 2007. *L'acquisizione del linguaggio*. Milano: Cortina.

Guasti, M. T. 2004. *Language Acquisition. The growth of grammar*. Cambridge, MIT Press.

Guasti, M. T., de Lange, J., Gavarró, A. and Caprin, C. 2003. Article Omission: Across Child Languages and across Special Registers. Van Kempen, J. and Baauw, S. (eds.) *Proceedings of GALA 2003 Generative Approaches to Language Acquisition*, Volume 1.

Guasti, M. T., de Lange, J., Gavarró, A. and Caprin, C. 2008. Article omission: across child languages. *Language Acquisition*. 15, pp89–119.

Hinzelin, O. The acquisition of subjects in bilingual children. Pronoun use in Portuguese-German children. In Natascha Müller (ed.). *(In)vulnerable domains in multilingualism*. John Benjamins. pp107-137.

Hirsh-Pasek, K., and Golinkoff, R. M. 1996. The preferential looking paradigm reveals emerging language comprehension. In McDaniel, D., McKee, C. and Cairns, H. (Eds.), *Methods for assessing children's syntax*. Cambridge, MA: MIT Press. pp105-124.

Hoff, E. and Naigles, L. 2002. How children use input to acquire a lexicon. *Child Development*. 73 (2), pp418-433.

Hoff-Ginsberg, E. 1991. Maternal speech and the child's development of syntax: a further look. *Journal of Child Language*. 17 (1), pp85-99.

Holmberg, A. and Roberts, I. (eds.) 2009. *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambidge University Press.

Hopp, H. 2007. *Ultimate attainment at the interfaces in second language acquisition: grammar and processing*. PhD thesis. Rijksuniversiteit Groningen.

Hulk, A. and Müller, N. 2000. Crosslinguistic influence at the interface between syntax and pragmatics. *Bilingualism: Language and Cognition*. 3 (3), pp227-244.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., Levine, S. 2002. Language input and child syntax. *Cognitive Psychology*. 45 (3), pp337-374.

Hyams, N. 1986. *Language acquisition and the theory of parameters*. Dordrecht: Reidel Publishing Company.

Jaeggli, O. and Safir, K.J. (eds.) 1989. *The null subject parameter*. Springer.

Juan-Garau, M. and Pérez-Vidal, C. 2001. Mixing and pragmatic parental strategies in early bilingual acquisition. *Journal of Child Language*. 28 (1), pp59-86.

Jurugo, D. A. 2009. Dialogue in the family. How often do adults engage children in dialogue? Paper presented at *The Child: Postgraduate Research Conference,* Belfast, 2-3 June 2009.

Kapetangianni, K. 2007. The development of peripheral positions in early child grammars. In Caunt-Nulton, H. Kulatilake, S. and Woo, I-H. (eds.), *Proceedings of the 31st Boston University Conference on Language Development vol. 2,* pp358 - 369.

Kapetangianni, K. 2008. The development of syntax-pragmatics connections in early child grammars: evidence from Greek, *Proceedings* of the *42nd Annual Meeting of the Chicago Linguistic Society*.

Klammer, A. 2006. *Erwerb und Beschaffenheit des deutsch-italienisch bilingualen Lexicons in der holophrastischen Phase. Eine quantitative und qualitative Analyse*. Ms dissertation. University of Graz.

Koppe, R. 1996. Language differentiation in bilingual children: the development of grammatical and pragmatic competence. *Linguistics*. 34, 5(345), pp927-954.

Kupisch, T. 2007. Determiners in bilingual German–Italian children: What they tell us about the relation between language influence and language dominance. *Bilingualism: Language and Cognition*. 10 (1), pp57-78.

Kupisch, T. and Bernardini, P. 2007. Determiner use in Italian-Swedish and Italian-German children: do Swedish and German represent the same Parameter Setting? *Nordlyd*. 34 (3), pp209-229.

Lanza, E. 1992. Can bilingual two-year-olds code-switch? *Journal of Child Language*. 19, pp633-658.

Lanza, E. 1997. *Language mixing in infant bilingualism: a sociolinguistic perspective*. Oxford: Clarendon Press.

Lidz, J., Gleitman, H. and Gleitman, L. 2003. Understanding how input matters: verb learning and the footprint of Universal Grammar. Cognition, 2003. 87 (3), pp151-178.

Long, M. H. 1990. Maturational constraints on language development. *Studies in Second Language Acquisition*. 12 (3), pp251-285.

López, L. 2009. *A derivational syntax for information structure*. Oxford University Press.

Lorusso, P., Caprin, C. and Guasti, M. T. 2004. Overt subject distribution in early Italian children. In *Boston University Conference on Language Development*, Boston.

Lozano, C. 2009. Selective deficits at the syntax-discourse interface: Evidence from the CEDL2 corpus. In Snape, N., Leung, Y. and Sharwood-Smith, M. (eds). *Representational deficits in SLA*. Amsterdam: John Benjamins, pp127–166.

McLaughlin, B. 1984. Early bilingualism: methodological and theoretical issues. *Neurolinguistics*. 13, pp19-45.

Meisel, J.M. 2001. The simultaneous acquisition of two first languages: Early differentiation and subsequent development of grammars. In Cenoz, J. and Genesee, F. (eds.) *Trends in Bilingual Acquisition*. pp11-41.

Meisel, J.M. 2004. The bilingual child. In Bhatia, T. K. and Ritchie, W. C., *The Handbook of bilingualism*. Blackwell Publishing. pp91-113.

Meisel, J.M. 2006. The development of the weaker language in bilingual first language acquisition. *Paper presented at Language acquisition and bilingualism: consequences for a multilingual society*, Toronto, 4-7 May 2006.

Meisel, J.M. 2007. The weaker language in early child bilingualism: acquiring a first language as a second language? *Applied Psycholinguistics*. 28 (3), pp495-514.

Meisel, J.M. 1986. Word Order and Case Marking in Early Child Language. Evidence from Simultaneous Acquisition of Two First Languages: French and German. *Linguistics*. 24, pp123-183.

Meisel, J.M. 1994. Code-switching in young bilingual children: the acquisition of grammatical constraints. *Studies in Second Language Acquisition*. 16 (4), pp413-439.

Meisel. J.M. 2008. Morphological development in early child second language acquisition. *Paper presented at the XI Congress of the International Association for the Study of Child Language*. Edinburgh, 28 July-1 August 2008.

Montrul, S. 2004. Subject and object expression in Spanish heritage speakers: a case of morphosyntactic convergence. *Bilingualism: Language and Cognition*. 7 (2), pp125–142.

Müller, N, Cantone, K. Kupisch, T. and Schmitz, K. 2002. On the language influence in bilingual first language acquisition: Italian-German, *Linguistische Berichte*. 190, pp157-206.

Müller, N. 2008. Bilingual first language acquisition at the interface between syntax and pragmatics. *EUROSLA Yearbook 8*. John Benjamin. pp52-78.

Müller, N. and Hulk, A. 2001. Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4 (1), 1-21.

Müller, N. and Pillunat, A. 2008. Balanced bilingual children with two weak languages: a French/German case study. In Guijarro-Fuentes, P., Larrañaga, M. P and Clibbens, J. (eds.), *First language acquisition of morphology and syntax. Perspectives across languages and learners*. John Benjamin. pp269-294.

Nicoladis, E. 1998. First clues to the existence of two input languages: pragmatic and lexical differentiation in a bilingual child. *Bilingualism: Language and Cognition*. 1 (2), pp105-116.

Nicoladis, E. 2008. Bilingualism and language cognitive development. In Altarriba, J. and Heredia, R.R. (eds.) An introduction to bilingualism: principles and processes. CRC press. pp167-181.

O'Grady, W. 1997. *Syntactic development*. University of Chicago Press.

Oller, D.K., Eilers, R.E., Urbano, R. and Cobo-Lewis A. B. 1997. Development of precursors to speech in infants exposed to two languages. *Journal of Child Language*. 24, pp407-425.

Padilla, A.M. and Liebman E. 1975. Language acquisition in the bilingual child. *The Bilingual Review/La Revista Bilingue*. 2 (1-2), pp34-55.

Pan, B.A, Rowe, M.L., Singer, J. and Snow, C. 2005. Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*. 76 (4), pp763-782.

Pancsofar, N. and Vernon-Feagans, L. 2006 Mother and father language input to young children: contributions to later language development, *Journal of Applied Developmental Psychology*. 27 (6), pp571–587.

Paradis, J. and Navarro, S. 2003. Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: what is the role of the input? *Journal of Child Language*. 30 (2), pp371-393.

Petersen, J. 1988. Word-internal code-switching constraints in a bilingual child's grammar. *Linguistics*. 26, pp479-493.

Pinto, M. 1997. *Licensing and interpretation of inverted subjects in Italian*. Ph.D. dissertation. Utrecht Institute of Linguistics OTS, Utrecht.

Polinski, M. 1997. American Russian: Language loss meets language acquisition. In Browne, W. et al. (eds.), *Formal approaches to Slavic linguistics*, Ann Arbor: Michigan Slavic Publications. pp370-407.

Polinsky, M. 2008. Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal*. 6 (1), pp40-71.

Radford, A. 1990. *Syntactic theory and the acquisition of English syntax*. Oxford: Blackwell.

Rangone, M. and Sgaggio, F. 2007. *Italiani in movimento. Progetto di ricerca sulla recente migrazione italiana in Irlanda.* Unpublished manuscript.

Redlinger, W. and Park, T.-Z. 1980. Language Mixing in Young Bilinguals. *Journal of Child Language*. 7 (2), pp337-352.

Ritchie, W.C. and Bhatia, T.K. 2004. Social and psychological factors in language mixing. In Bhatia, T.K. and Ritchie, W.C. (eds.) *The handbook of bilingualism*. Blackwell. pp336-352.

Rizzi, L. 1982. *Issues in Italian syntax*. Foris Publications, Dordrecht.

Rizzi, L. 1994. Root null subjects and early null subjects. In Hoekstra, T. e Schwartz, B. (eds.). *Language acquisition studies in generative grammar.* Amsterdam: John Benjamins. pp151-176.

Rizzi, L. 1994. Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition*. 3, pp371-393.

Rolla San Francisco, A., Maria, C., August, D. and Snow, C. E. 2006. The role of instruction and vocabulary in the English phonological awareness of Spanish-English bilingual children. *Applied Psycholinguistics*. 27, pp.229-246.

Rothman, J. 2007. Heritage speaker competence differences, language change, and input type: inflected infinitives in Heritage Brazilian Portuguese. *International Journal of Bilingualism*. 11 (4), pp359-389.

Rothman, J. 2009. Pragmatic deficits with syntactic consequences?: L2 pronominal subjects and the syntax-pragmatics interface. *Journal of Pragmatics*. 41 (5), pp951-973.

Rothweiler, M. 2008. Verbal inflection, case morphology and the acquisition of sentence structure in early successive bilinguals. *Paper presented at the XI Congress of the International Association for the Study of Child Language*. Edinburgh, 28 July-1 August 2008.

Schlyter, S. 1993. The weaker language in bilingual Swedish-French children. In K. Hyltenstam & A. Vilberg (eds.), *Progression and regression in language*. Cambridge: Cambridge University. pp289-308.

Schlyter, S. 1994. Early morphology in Swedish as the weaker language in French-Swedish bilingual children. *Scandinavian Working Papers on Bilingualism*, 9. University of Stockholm.

Schmitz, K. and Muller, N. 2008. Strong and clitic pronouns in monolingual and bilingual acquisition of French and Italian. *Bilingualism: Language and Cognition*. 11 (1), pp19-41.

Serratrice, L. 1999. *The emergence of functional categories in bilingual first language acquisition*. Ph.D thesis, University of Edinburgh.

Serratrice, L. 2005. The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied Psycholinguistics*. 26, pp437–462.

Serratrice, L., Sorace, A. and Paoli, S. 2008. Cross-linguistic influence at the syntax-pragmatics interface: subjects and objects in Italian-English

bilingual and monolingual acquisition. *Bilingualism Language and Cognition*. 7 (3), pp183-206.

Sorace, A. 2003. Near-Nativeness. Doughty, C. J. and Long, M. H. (eds.) *The Handbook of Second Language Acquisition*. Blackwell Publishing. pp130-151.

Sorace, A. and Serratrice, L. 2009. Internal and external interfaces in bilingual language development: beyond structural overlap. *International Journal of Bilingualism* 13 (2), pp195-210.

Sorace, A., Serratrice, L. Filiaci, F., Baldo, M. 2009. Discourse conditions on *subject* pronoun realization: testing the linguistic intuitions of older bilingual children. *Lingua* 119 (3), pp460-477.

Snow, C. E. 1972. Mothers' speech to children learning language. *Child Development*. 43, pp549-565.

Stoll, S. and Lieven, E. 2008. Comparing young children's communicative environments in two cultures. *Paper presented at the XI Congress of the International Association for the Study of Child Language*. Edinburgh, 28 July-1 August 2008.

Strömqvist, S. and Richthoff, U. 1999. Linguistic feedback, input and analysis in early language development. *Journal of Pragmatics*. 31 (10), pp1245-1262.

Taeschner, T. 1983. *The sun is feminine : a study on language acquisition in bilingual children*. Springer.

Taeschner, T. and Volterra, V. 1978. The acquisition and development of Language by bilingual children. *Journal of Child Language*. 5 (2), pp311-326.

Theakston A.L., Lieven E., Pine J., Rowland C. 2004. Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language*, 31. pp61-99.

Tomasello M. and Brooks, P. 1999. Early syntactic development. In: Barrett M. (ed.), *The development of language*. UCL Press. pp161-190.

Unsworth, S. 2005. On the syntax-semantics interface in Dutch: Adult and child L2 acquisition compared. *International Review of Applied Linguistics*. 42 (4), pp173-187.

Unsworth, S. 2007. L1 and L2 acquisition between sentence and discourse: Comparing production and comprehension in child Dutch. *Lingua*. 117 (11), pp1930-1958.

Unsworth, S. 2008. Comparing child L2 development with adult L2 development: How to measure L2 proficiency. In Gavruseva, E. and Haznedar, B. (eds.) *Current Trends in Child Second Language Acquisition*. Amsterdam: John Benjamins. pp301-336.

Valian, V. and Casey, L. 2003. Young children's acquisition of wh-questions: The role of structured input. *Journal of Child Language*. 30, pp117-143.

Vallauri, E. 2004. The relation between mind and language: The Innateness Hypothesis and the Poverty of the Stimulus. *The Linguistic Review*. 21, pp345-387.

Vigil, D. C., Hodges, J., and Klee, T. 2005. Quantity and quality of parental language input to late-talking toddlers during play. *Child Language Teaching and Therapy*. 21, pp107-122.

Vihman, M.M. 1985. Language differentiation in the Bilingual Infant. *Journal of Child Language*. 12 (2), pp297-324.

Wei, L. 2000. Dimensions of bilingualism. In *The bilingualism reader*. Routledge. pp3-25.

Westergaard. M. R. 2004. The interaction of input and UG in the acquisition of verb movement in a dialect of norwegian. *Nordlyd: Tromsø Working Papers in Language Acquisition*. 32 (1), pp110-134.

White, L. and Genesee, F. 1996. How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*. 12 (3), pp233-265.

Wijnen, F., Kempen, M. and Gillis, S. 2001. Root infinitives in Dutch early child language: an effect of input? *Journal of Child Language.* 28, pp629–660.

Wilson, F. 2009. *Processing at the syntax discourse interface in second language acquisition.* PhD. University of Edinburgh.

Windsor, J., Glaze, L.E. and Koga, S.F. 2007. Language acquisition with limited input: Romanian institution and foster care. *Journal of Speech, Language, and Hearing Research*. 50 (5), pp1365-1381.